Understanding Allosteric Enzyme Regulation Using Macromolecular Rate Theory

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science (Research) in Biological Sciences

at

The University of Waikato

by

Carlin James Hamill

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2020

Abstract

Enzyme catalysed reaction rates produce a curved temperature dependence with a temperature optimum, T_{opt}. Traditionally enzyme temperature dependence has been modelled by the Arrhenius and Eyring-Polanyi equations. Deviations from these models at elevated temperatures have traditionally been attributed to enzyme denaturation. Increasing evidence shows that enzyme denaturation is insufficient to fully explain the curvature for enzyme catalysed rates with temperature.

Macromolecular rate theory (MMRT) accounts for these deviations from early models by introducing the concept of heat capacity (C_P). During an enzymatic reaction, the heat capacity is reduced as the transition state complex forms (due to enzyme rigidification), resulting in an overall negative change in heat capacity ($\Delta C_P^{\ddagger}$) for the reaction. Heat capacity is a function of the number and energies of the vibrational modes for the individual molecules. Enzyme-ligand binding will influence the energies of the vibrational modes of the molecular complex therefore affecting an enzymes' heat capacity and its catalytic rate via $\Delta C_P^{\ddagger}$. Notably, a form of enzyme regulation (called allostery) utilises ligand binding at a site other than the active site to effect changes in catalytic rate.

The underlying mechanisms of allostery are poorly understood. Here, the mechanisms of allostery are explored using a model isomaltase enzyme and MMRT to investigate changes in heat capacity. Several mutants of the model MalL enzyme were designed, produced and purified. These mutations aimed to either alter the structural integrity or mimic ligand binding in the C-terminal domain. Using a Stopped-Flow apparatus, these mutants were characterised kinetically and using crystallographic techniques, two structures were solved at high resolution.

Targeted mutations approximately 30 Å from the active site showed significant changes in enzyme rate. Overall, the two new structures of MalL were very similar to the apo-enzyme showing only very minor structural deviations. Changes to the rate can thus be attributed to changes in enzyme dynamics and heat capacity. These results are a promising first step in linking changes in enzymatic activity with changes to enzyme dynamics and heat capacity, using MMRT as a model.

Acknowledgements

Firstly, thanks to my primary supervisors. Thanks to Professor Vic Arcus for your guidance and endless enthusiasm about all things MMRT. Thanks for always having a helpful suggestion when things are going wrong and pointing me in the right direction. Thanks to Dr Emma Summers for providing me with the skills and knowledge to successfully complete this project, and for your helpful critiques of this thesis. Your help has been invaluable along this journey. Thanks also to Dr Erica Prentice for helping me get my start in the lab.

Secondly, thanks to Dr Claire Mulholland and Dr Scott Cameron for your help with teaching me the essentials of crystallography.

Thirdly, thanks to Andrew Howard, Emily Grout, Keely Oldham and everybody else in and around C.2.10 for making the lab such a great place to work. Your support, encouragement and friendliness mean a lot to me and has made this journey enjoyable even when things were going wrong in the lab. A special thanks to Dr Judith Burrows for your positivity and help. This lab would not work nearly so well without you.

Thanks to the University of Waikato for their support through the University of Waikato Research Masters Scholarship.

Finally, thanks to my family, especially Mum and Dad for supporting and encouraging me throughout my studies and for listening to me trying to explain ‘enzymes’. Doing this wouldn’t have been possible without your support.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

A:	Temperature dependence factor of $\Delta C_P^{\ddagger}$
AMP:	Ampicillin
APS:	Ammonium persulphate
ATP:	Adenosine triphosphate
Cp:	Heat capacity at constant pressure
DNA:	Deoxyribonucleic acid
E_A:	Activation energy of a chemical reaction
EDTA:	Ethylenediaminetetraacetic acid
FPLC:	Fast protein liquid chromatography
G:	Gibbs free energy
H:	Enthalpy
IMAC:	Immobilised metal affinity chromatography
IPTG:	Isopropyl β-D-1-thiogalactopyranoside
KNF:	Koshland, Nemethy, Filmer allosteric model
LB:	Luria broth
MD:	Molecular dynamics
MMRT:	Macromolecular rate theory
MQ:	Milli-Q ultrapure water
MWC:	Monod, Wyman, Changeaux allosteric model
nm:	Nanometre
OD:	Optical density
PAGE:	Polyacrylamide gel electrophoresis
PCR:	Polymerase chain reaction
PNG:	*p*-Nitrophenyl-α-D-glucopyranoside
R:	Ideal gas constant
RMSD:	Root Mean Square Deviation
S:	Entropy
SDS:	Sodium dodecyl sulphate
SEC:	Size exclusion chromatography

T:	Temperature
T_0:	Reference temperature for MMRT temperature dependant $\Delta C_P^{\ddagger}$
T_{opt}:	Temperature optimum for enzyme catalysed rate
TEMED:	Tetramethylethylenediamine
Tris:	2-Amino-2-(hydroxymethyl)propane-1,3-diol
TST:	Transition state theory
UV:	Ultraviolet
WT:	Wildtype

1 Introduction

1.1 Enzyme Catalysis

Enzymes have an extraordinary ability to catalyse chemical reactions such that they occur with dramatic rate improvements compared to the uncatalysed reaction. Enzymatic reactions can have rate improvements of up to 10^{17} times (Fersht, 1999). Understanding the mechanisms underlying these rate increases is therefore of considerable interest. In general, enzymes are considered to lower the activation energy of a reaction compared to the uncatalysed reaction. The activation energy (shown as $\Delta G^{\ddagger}$ in Figure 1.1) is the difference in Gibbs free energy (ΔG) between the reactants (enzyme and substrate) (ES) and the transition state complex ($ES^{\ddagger}$). The Gibbs free energy (G) is a measure of the available energy in a system. A lower activation energy increases the rate of reaction due to an increased ability of the substrates to cross the activation energy barrier. The transition state is a high energy pseudo-intermediate that is formed during the conversion of substrate to the end product (Garrett & Grisham, 2013).

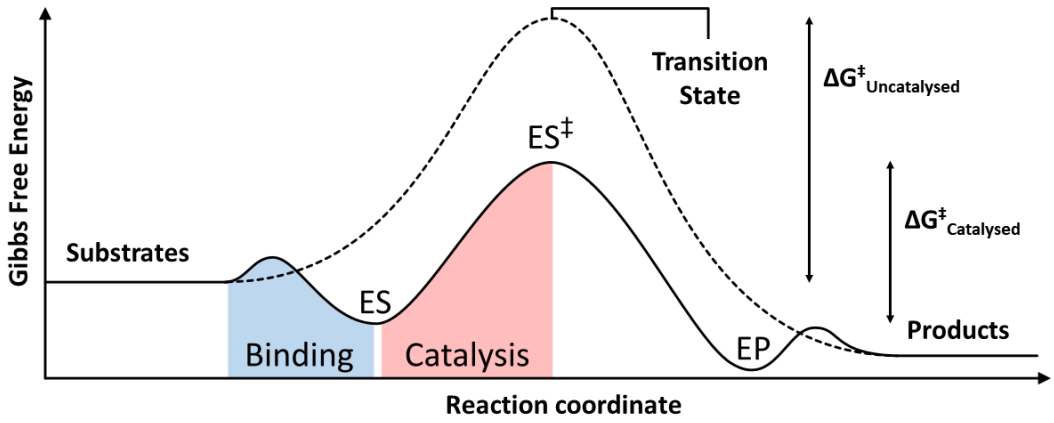

Figure 1.1. Gibbs free energy changes along an enzyme-substrate reaction coordinate. Dashed line shows uncatalysed reaction with activation energy given by $\Delta G^{\ddagger}_{Uncatalysed}$. Solid line shows catalysed reaction with stabilisation of transition state ($ES^{\ddagger}$) and lowered activation energy given by $\Delta G^{\ddagger}_{Catalysed}$. Adapted from Shafee (2015).

Enzymes use various methods to achieve this decrease in activation energy. Enzymes are highly specific for their substrates and it is this specificity which first gave insight into how enzymes may achieve their rate improvements. Strict enzyme specificity lead to the development of hypotheses on how enzymes interact with their substrates (Berg *et al.*, 2015). The first model was the "lock-and-key" hypothesis, where the active site of the unbound enzyme is complementary in shape to the substrate, allowing the substrate to bind to the active site, which results in activation of the substrate (Fischer & Pinner, 1894). This however failed to provide any insight to the basis of an enzyme's catalytic proficiency. Substrate binding was then proposed to bind in an "induced fit" mechanism, where the enzyme changes shape upon substrate binding, inducing a certain strain on it, introducing the first ideas on the mechanisms of enzyme catalysis (Haldane, 1930).

The idea of induced fit was expanded with the hypothesis that the transition-state is stabilised by the enzyme – postulated by Pauling (1948), which lead to a new understanding of enzyme catalysis. Pauling's understanding of how a molecular crystal structure developed by way of the complementarity of each molecule in the structure (excluding all other molecules that did not precisely fit) lead to his theory that the same mechanism is responsible for biological specificity (Pauling, 1948). He proposed that enzymes are complementary in structure to the complexes of reactions that they catalyse, in other words, the transition state. His theory was that the enzyme has a detailed complementariness in molecular structure for the transition state which leads to a decrease in the energy of the complex of the two. This in turn would lead to a decrease in the energy of activation of the reaction, hence increasing the rate of the reaction (Pauling, 1948).

Evidence for the induced fit and transition state stabilisation was provided by the crystal structure of lysozyme. The lysozyme structure was published in 1965 and was the first enzyme to have its structure solved (Blake *et al.*, 1965). Following this, efforts turned to getting a crystal structure of lysozyme bound to a substrate (Phillips, 1966). Lysozyme catalyses the degradation of large chain polysaccharides such as peptidoglycan which consists of alternating units of *N*-acetylglucosamine

(NAG) and *N*-acetylmuramic acid (NAM). The structure of lysozyme was able to be solved with the trisaccharide $(NAG)_3$ bound in the top portion of the active site cleft. It had been noted that NAG itself as well as the trisaccharide acted as an inhibitor of lysozyme. It therefore seemed likely that $(NAG)_3$ bound as a competitive inhibitor in the same way as the substrate. This structure revealed the molecular details of the enzyme's interaction with the substrate. The structure to the left of the active site cleft was found to shift to close the active site cleft when substrate is bound (Figure 1.2). In the substrate bound structure tryptophan residue 62 moves 0.75 Å to encompass the bound substrate (Phillips, 1966). This structural change upon ligand binding supported the induced fit model of enzyme-substrate interactions.

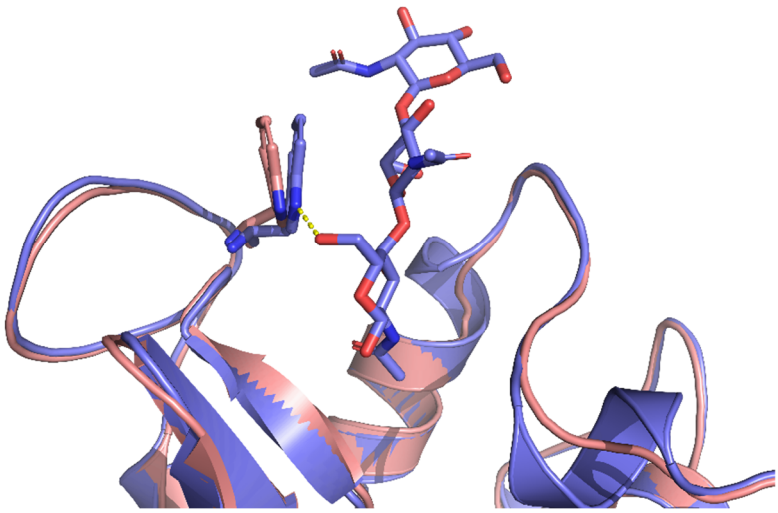

Figure 1.2. Crystal structure of lysozyme with (Blue; PDB: 1HEW) and without (Red; PDB: 1LYZ) substrate $(NAG)_3$ bound showing structural changes upon substrate binding.

It was noted that the $(NAG)_3$ substrate only filled half of the active site cleft. Additional sugar molecules were able to be modelled into the active site (Phillips, 1966). For additional sugar units to bind, the fourth sugar unit was required to be bound in a distorted conformation. Lysozyme was found to cleave the glycosidic bond between the fourth and fifth sugar units. Two reactive amino acids (Asp^{52} and Glu^{35}) were identified in the vicinity of this bond. From this a mechanism for the catalytic activity of lysozyme was able to be proposed.

The cleavage of glycosidic bonds involves the formation of a carbenium ion intermediate. Lemieux and Huber showed in 1955 that when NAG incorporates a carbenium ion at carbon-1 (as it does during the lysozyme reaction) in the absence of enzymatic interactions it also adopts the same strained conformation shown by the fourth lysozyme sugar unit (Lemieux & Huber, 1955). This suggests that substrate binding to lysozyme favours the conformation that promotes formation of the carbenium ion required for catalysis (Phillips, 1966). The carbenium intermediate formed during the reaction is further stabilised by the two amino acids Asp^{52} and Glu^{35} (Benkovic & Hammes-Schiffer, 2003) (Figure 1.3). The interaction of residues implicit with substrate binding and hydrolysis provided evidence for the theory that the enzyme bound preferentially to the transition state (Phillips, 1966).

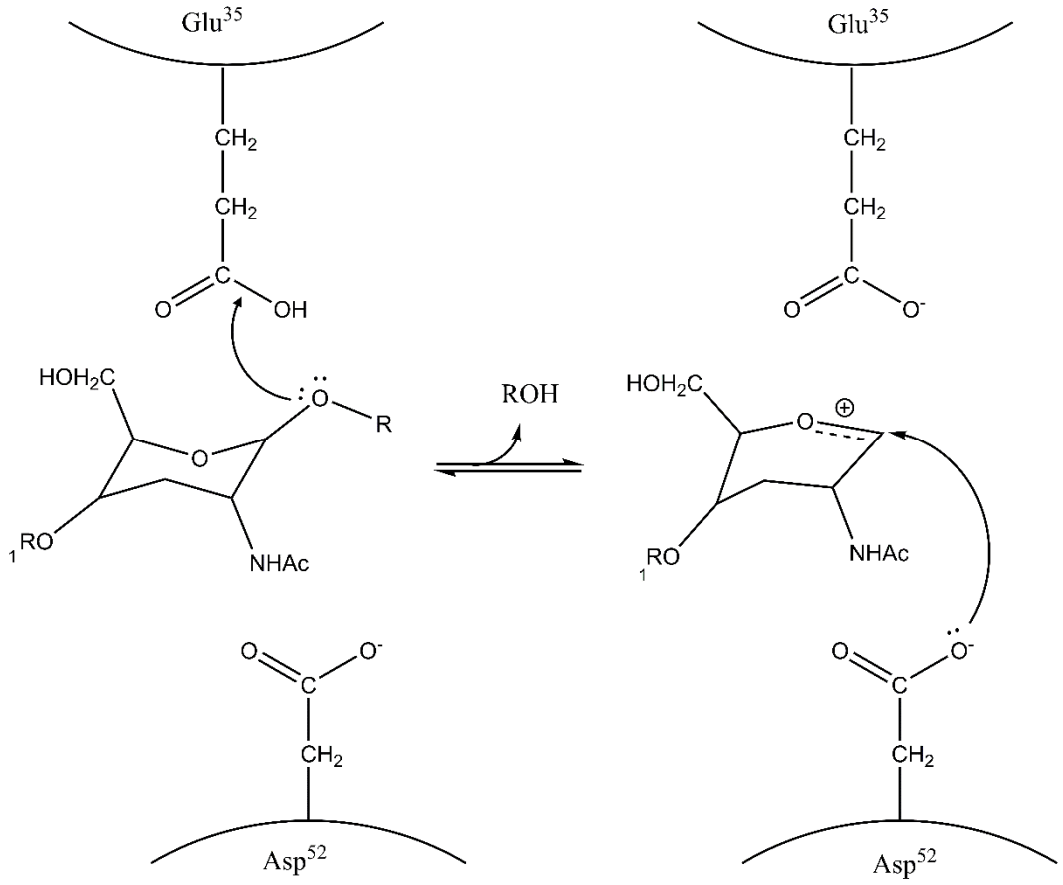

Figure 1.3. N-acetylglucosamine substrate bound to lysozyme provides evidence for transition state stabilisation. Amino acids Asp^{52} and Glu^{35} are positioned to stabilise the formation of oxocarbenium intermediate. Adapted from Strynadka and James (1991).

1.1.1 Chemical Catalysis

One of the ways enzymes catalyse reactions is by providing relevant chemical moieties in the correct orientations to provide chemical action to catalyse a reaction. There are three general mechanisms of enzyme chemical catalysis (Berg *et al.*, 2015).

1.1.1.1 Covalent catalysis

Covalent catalysis utilises a reactive group that briefly becomes covalently attached to the substrate during catalysis. Covalent bonds involve the sharing of electron pairs. The reactive group often acts as a powerful nucleophile and participates in nucleophillic attack on the substrate to facilitate a reaction. Nucleophiles have excess electron density and can donate electron pairs to form a bond. For example, peptide proteases utilise covalent catalysis to catalyse the cleavage of peptide bonds by hydrolysis (Berg *et al.*, 2015). During peptide hydrolysis water acts as a nucleophile to facilitate the cleavage of the peptide bond (Figure 1.4).

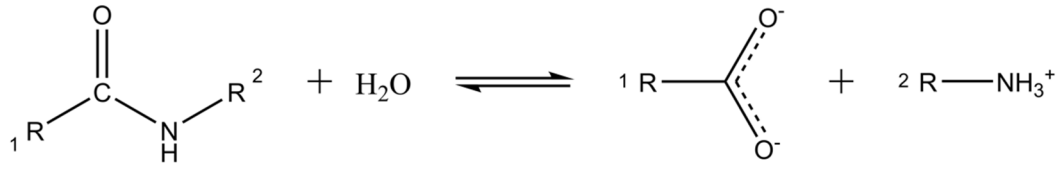

Figure 1.4. Covalent catalysis is illustrated by peptide bond hydrolysis, which is facilitated by the addition of water (acting as a nucleophile) across the peptide bond. Adapted from Berg et al. (2015).

The hydrolysis of a peptide bond is thermodynamically favoured but kinetically unfavoured. During hydrolysis the resonance of the peptide bond renders it resistant to hydrolysis (Berg *et al.*, 2015). This is due to the sharing of electron density from the carbonyl group into the carbon-nitrogen bond by resonance, giving it a partial double bond character, which strengthens the bond and renders it resistant to hydrolysis (Figure 1.5). The carbonyl group is also less electrophilic (ability to accept electrons) and resists nucleophillic attack (Berg *et al.*, 2015). An

enzyme must overcome this disfavoured kinetics and facilitate a nucleophillic attack on the carbonyl group to successfully facilitate peptide bond hydrolysis.

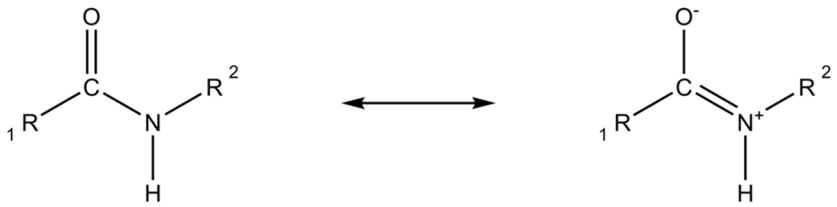

Figure 1.5. Resonance of the peptide bond gives the carbon-nitrogen bond a partial double bond character, strengthening it and making it more difficult to hydrolyse. Adapted from Berg et al. (2015).

A good example of this mechanism is the protease chymotrypsin. Chymotrypsin catalyses the cleavage of peptide bonds on the carboxyl side of large hydrophobic residues. It utilises a catalytic serine residue to facilitate a nucleophillic attack on the carbonyl group. During catalysis the serine residue is deprotonated by a nearby histidine residue forming a strongly nucleophillic alkoxide ion which attacks the peptide carbonyl to form a covalently bonded acyl-enzyme intermediate (Berg *et al.*, 2015). The addition of water then allows for the release of the hydrolysed substrate and regeneration of the active site.

1.1.1.2 Acid-Base catalysis

In acid-base catalysis an enzyme residue plays the role of proton donor or acceptor in place of water. Acid base catalysis can stabilise charges on reaction intermediates by transfer of protons (Fersht, 1999). In general-base catalysis positive charges are stabilised by proton transfer to a base. Similarly, acid catalysis stabilises negative charges by donating a proton (Fersht, 1999).

Alpha amylases are a group of enzymes that utilise an acid-base mechanism to catalyse the cleavage of glycosidic bonds (Davies & Henrissat, 1995). In this reaction the glycosidic oxygen is protonated by an acid group (A) with nucleophilic assistance from a base group (B), resulting in the formation of a glycosyl enzyme intermediate (Figure 1.6). This intermediate is hydrolysed by water producing the product and regenerating the active site.

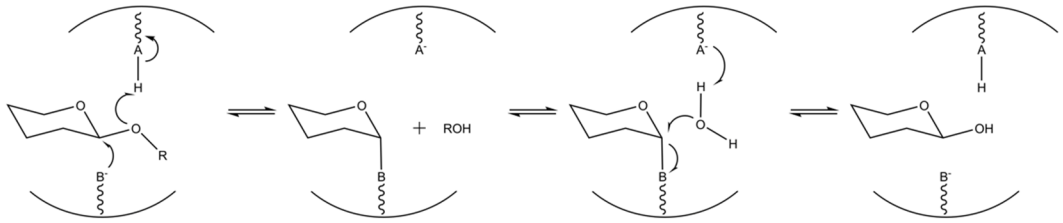

Figure 1.6. Acid Base catalysis of glycosidic cleavage by a retaining alpha-amylase. Glycosidic oxygen is protonated by acid group (A) with nucleophillic assistance from base group (B) forming a glycosyl enzyme intermediate. Intermediate is hydrolysed by water completing the reaction and regenerating the active site. Adapted from Davies and Henrissat (1995).

1.1.1.3 Metal-Ion catalysis

Metal ions can play several roles in catalysis (Berg *et al.*, 2015). Metal ions can assist in the formation of nucleophiles by direct coordination. For example, zinc is used in carbonic anhydrase catalysis to stabilise the formation of the negatively charged hydroxide nucleophile utilised in the reaction (Figure 1.7). Metals can also stabilise negative charges in transition state intermediates by acting as an electrophile. This reduces the free energy of the intermediate which lowers the activation energy required for the reaction to occur. Metal ions can also assist in substrate binding and help to bind the substrate in the correct orientation for catalysis. Another important metal ion is magnesium. For example, many ATP binding enzymes use magnesium to bind ATP (Maguire & Cowan, 2002). The magnesium ion has a large hydration sphere so can help to position water correctly for participation in the enzymatic mechanism (Maguire & Cowan, 2002).

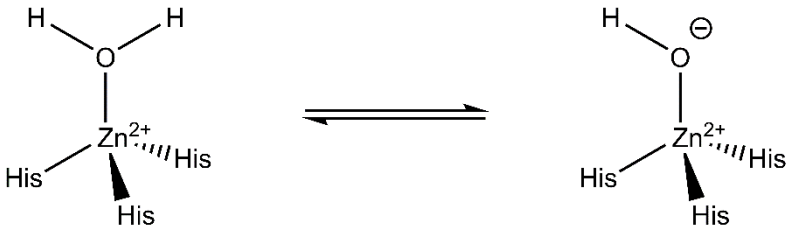

Figure 1.7. Zinc assists in the formation of hydroxide nucleophile utilised in the catalytic mechanism of carbonic anhydrase. Adapted from Berg et al. (2015).

1.1.2 Dynamic mechanisms of enzyme rate acceleration

1.1.2.1 Ground state destabilisation

Rate improvements in enzyme catalysis can occur by either raising the energy of the enzyme bound to substrate (ES), known as the ground state or lowering the energy of the transition state complex (TS), either of which will lower the activation energy required for a reaction (Figure 1.8).

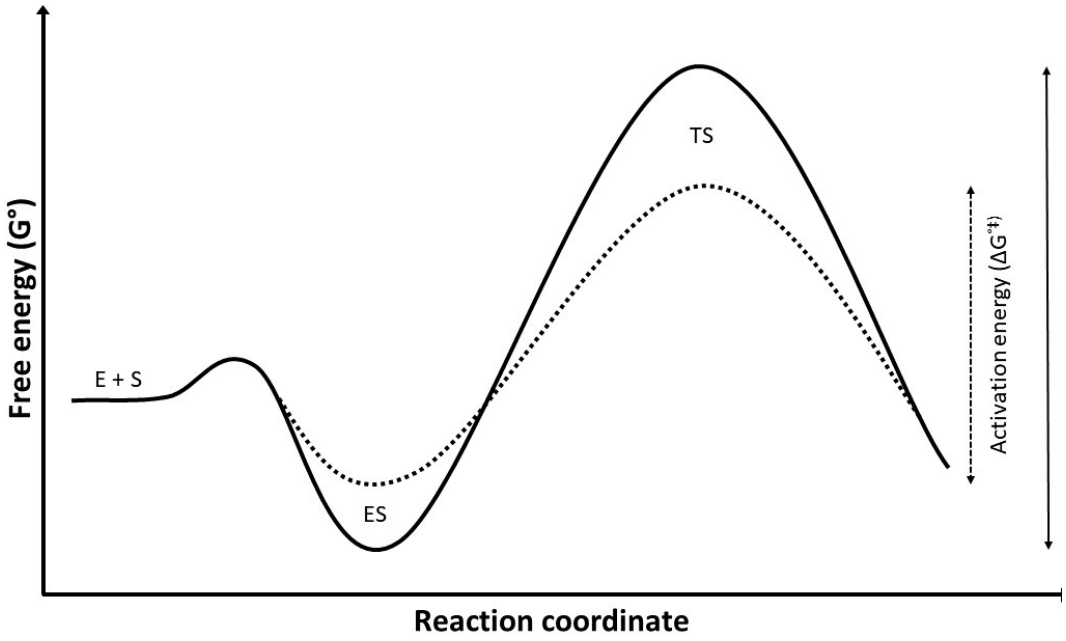

Figure 1.8. Enzyme rate improvements can be achieved by lowered activation energy ($\Delta G^{o\ddagger}$). Dashed line shows increased energy of ground state (ES) and lowered energy of transition state (TS) resulting in lowered activation energy (double arrow head dashed line).

Substrate binding to the enzyme results in an important catalytic compromise. There is a small release of Gibbs free energy during substrate binding ($\Delta G_{binding}$) (Berg *et al.*, 2015). This results in the ES complex having a lower energy state (or in other words a negative binding energy), however this ensures that binding of the substrate is favourable (Figure 1.9). This interaction should however not be too strong as a large $\Delta G_{binding}$ will increase the activation energy required to perform the reaction unreasonably and catalytic efficiency will be lost. This can be

countered by entropic effects of substrate binding to the enzyme leading to the destabilisation of the ground state (Fersht, 1999).

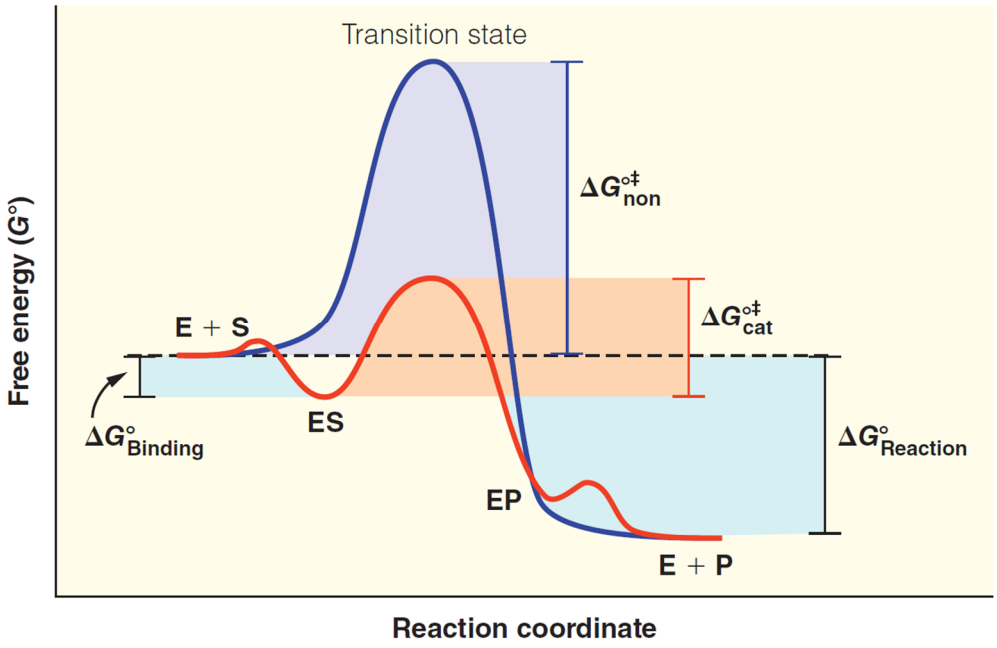

Figure 1.9. Effect of free energy release upon substrate binding on the activation energy required for enzyme catalysis. Blue line shows non-catalysed reaction with activation energy, $\Delta G^{\circ\ddagger}_{non}$. Red line shows enzyme catalysed reaction with activation energy, $\Delta G^{\circ\ddagger}_{cat}$. $\Delta G^{\circ\ddagger}_{cat}$ increases as the ground state (ES) is stabilised and $\Delta G^{\circ}_{binding}$ is increased (Appling et al., 2019).

Ground state destabilisation is one of the mechanisms enzymes use to raise the energy of the ES complex to ensure that the activation energy is not unnecessarily large due to excessive $\Delta G_{binding}$. Ground state destabilisation places a strain on the bound substrate to increase the energy of the ground state through electronic strain, bond distortion and conformational restriction (Berg *et al.*, 2015).

Electronic strain is an effect of the non-zero electric field present at the active site of an enzyme. The electric field generates a force on the substrate destabilising the ground state. The strength of this interaction depends on the strength of the electric field and the polarizability of the substrate. Tonge and Carey found that the energy involved in this interaction was almost completely reflected by a decrease in the activation energy (Tonge & Carey, 1992).

Bond distortion occurs when a substrate binds to an enzyme such that its bonds are either stretched or rotated. It was first suggested in 1930 that glycosidases bind their substrates in such a way that the glycosidic bond in strained (Haldane, 1930). This was later confirmed with the structure of the lysozyme and other glycosidic enzymes (Davies *et al.*, 1998; Phillips, 1966).

Ground state destabilisation can occur by a number of mechanisms. For example, substrate conformational restriction raises the free energy of the substrate by decreasing its entropy. In other words, in the enzyme substrate complex, there are fewer possible conformational states for the substrate and its entropy is decreased. For example, the glycoside hydrolase Cel5A from the bacteria *Bacillus agaradhaerens* binds its ribose substrate in a skew boat configuration. In this orientation the substrate is bound in a higher energy configuration, which contributes to the catalytic proficiency of the enzyme (Davies *et al.*, 1998).

1.1.2.2 Catalysis by Approximation

For reactions with two substrates, the catalytic rate can be enhanced by providing a site for the two substrates to interact in the correct geometry for a reaction to occur (Berg *et al.*, 2015). This increases the effective concentration of a substrate in solution. Binding of substrates confines the degrees of freedom (independent motion) and therefore lowers the entropy. This decrease in entropy by substrate binding lowers the overall entropy change for the reaction, thereby increasing the rate (Fersht, 1999).

1.2 Enzyme Kinetics

The most widely used model to understand enzyme kinetics is the Michaelis-Menten model developed by Leonor Michaelis and Maud Menten in 1913 (Fersht, 1999). Many enzymes follow Michaelis-Menten kinetics. Enzymatic reactions can be thought of as two parts, a reversible binding step, and a largely irreversible catalytic step (Chen *et al.*, 2010). This scheme produces Equation 1.1 where the rate of the substrate binding is given by k_1 and the rate of product formation is given by k_{cat}.

Equation 1.1. Michaelis-Menten reaction scheme

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P$$

The study of enzyme kinetics relies on the steady state approximation method. In this, it is assumed that the concentration of the enzyme substrate complex (ES) does not change (Berg *et al.*, 2015). This allows rates to be described with simplified equations. For the steady state approximation to hold, the kinetics are studied with excess substrate. Under these conditions the enzymes' rate is defined by the substrate concentration. As the substrate concentration increases the rate increases until it approaches the maximal rate (V_{max}) as shown in the rate versus substrate concentration curve below (Figure 1.10).

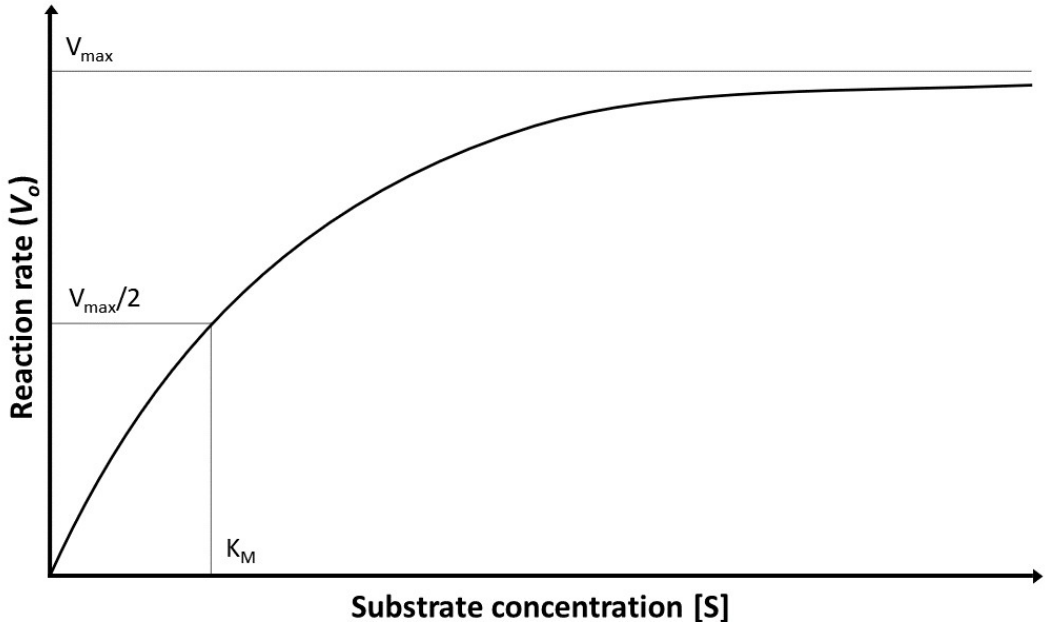

Figure 1.10. Michaelis-Menten Kinetics of an enzyme. Catalytic rate increases until enzyme is saturated and maximal rate (V_{max}) is reached. K_M is defined by the substrate concentration at half V_{max}.

An increase in rate as substrate concentration increases can be modelled by the Michaelis-Menten equation (Equation 1.2) which includes terms for the maximum rate and the Michaelis constant (K_M) (Berg *et al.*, 2015).

Equation 1.2. Michaelis-Menten equation

$$V_0 = \frac{V_{max}[S]}{K_M + [S]}$$

K_M is the equilibrium constant for substrate binding and is found on the graph in Figure 1.10 at the substrate concentration whereby the enzyme is performing at half of its maximal rate. It is also a variable that provides important information about the affinity of the enzyme for the substrate.

Another important measure of reactant rate is k_{cat} which can be calculated from V_{max} by Equation 1.3 (where $[E]_T$ represents total enzyme). k_{cat} provides a substrate independent measure of catalytic rate and represents the turnover of an enzyme per unit time.

Equation 1.3. Calculation of k_{cat}

$$V_{max} = k_{cat}[E]_T$$

1.3 Allostery

Allostery is the mechanism in which the effect of a ligand binding at a site distal to the active site, is transmitted through the enzyme allowing for the regulation of enzyme activity (Motlagh *et al.*, 2014). Allostery was first observed in the early 20th century by Christian Bohr who described the effect of carbon dioxide on oxygen binding to haemoglobin (Liu & Nussinov, 2016). Bohr observed that the affinity of haemoglobin for oxygen decreased as the partial pressure of carbon dioxide increased. The term allostery was coined by Jacques Monod and Francois Jacob in 1961 to describe enzyme inhibition where the inhibitor did not mimic the

substrate and exerted its effect by binding in its own distinct site away from the active site (Liu & Nussinov, 2016). In the following years two models were proposed to describe allostery, the KNF model by Koshland, Nemethy and Filmer (Koshland *et al.*, 1966), and the MWC model by Monod, Wyman and Changeaux (Monod *et al.*, 1965). Both the KNF and MWC models described allostery in terms of conformational changes, a concept that would dominate until recent years when a more dynamic view of allostery began to emerge (Liu & Nussinov, 2016). Despite these advances, the detailed mechanisms of allostery are not well understood.

1.3.1 Conformational models of allostery

The conformational models of allostery were heavily influenced by the allosteric effects observed in haemoglobin. The two main conformational models are the MWC and KNF models. Both of these models are phenomenological. That is, they both describe the action of allostery, without providing a suitable mechanism for the underlying changes (Cui & Karplus, 2008).

1.3.1.1 Monod-Wyman-Changeaux (MWC) Model (Symmetry model)

The core principle of the MWC model is that there is an equilibrium between two coexisting quaternary structural conformations (Eaton *et al.*, 2007). The MWC model requires a number of assumptions to be made. The allosteric protein must be an oligomer and exist in two states, a tense state (**T**) and a relaxed state (**R**) (Fersht, 1999). All subunits are assumed to have the same conformation and therefore the MWC model is also referred to as the symmetry model. The **T** state is more constrained than the **R** state and has a lower affinity for ligands. This lower affinity **T** state is more prevalent in the absence of ligands (Fersht, 1999). In each state equivalent binding sites on each monomer are identical and have equal binding affinities (Fersht, 1999). A central tenet of the MWC model is that both the **T** and **R** states are available independently of substrate or allosteric regulators being present. Allosteric regulators do not function by directly altering the structure of the **T** or **R** states, but by indirectly shifting the equilibrium between

these two states (Eaton *et al.*, 2007). In the MWC model allosteric enzymes are oligomeric with internal structural symmetry. The structural shift between the two states occurs by a rotation of monomer units. During this rotation symmetry in the oligomer is conserved (Cui & Karplus, 2008) (Figure 1.11).

The MWC model provided a simple explanation for allosteric control. It had been noted that many allosteric enzymes produced a sigmoidal rate vs substrate concentration curve in the absence of allosteric regulators (Figure 1.12). This is proposed to be the result of the substrate binding preferentially to the **R** state shifting the equilibrium to this state making it easier for additional substrate molecules to bind which produces a sigmoidal curve (Fersht, 1999). An allosteric activator functions by binding to the **R** state, increasing its concentration by shifting the equilibrium towards the **R** state. In some cases an allosteric activator may increase the prevalence of the **R** state to such an extent that the sigmoidal rate curve is shifted to show Michaelis-Menten kinetics (R-state curve) (Fersht, 1999). Whereas an allosteric inhibitor functions by binding to the **T** state and shifting the equilibrium towards this state, changing the rate curve (T-state curve) (Fersht, 1999).

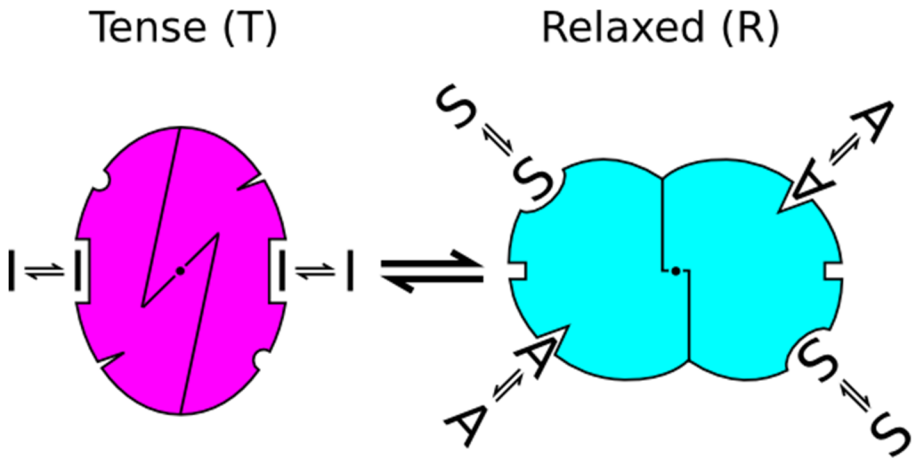

Figure 1.11. MWC model. Allostery is mediated by a conformational shift between the catalytically inactive tense state (T) and the active relaxed state (R), where substrate (S) can bind. Equilibrium between T and R can be shifted by binding of an inhibitor (I) or activator (A) (Le Novere, 2008).

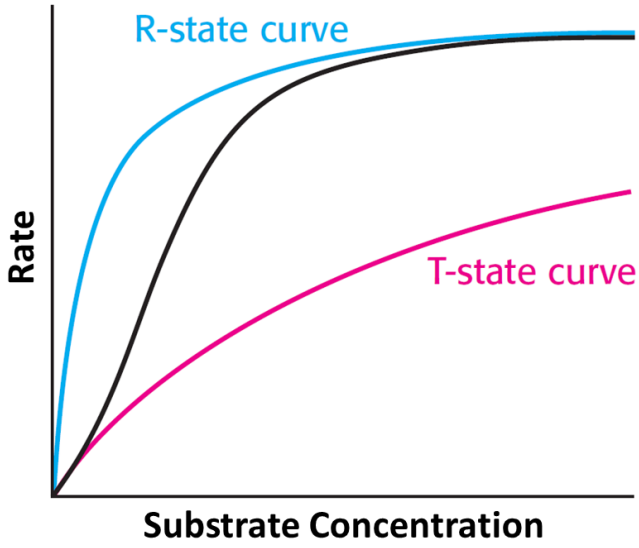

Figure 1.12. MWC Kinetics. An enzyme following MWC kinetics shows a sigmoidal (Black) rate curve. Addition of an inhibitor shifts rate curve to T-state kinetics (Pink). Addition of an activator shifts rate curve to R-state kinetics (Blue). Adapted from Berg et al. (2015).

1.3.1.2 Koshland-Nemethy-Filmer (KNF) Model (Sequential model)

Similar to the MWC model, the KNF model assumes there exists two states **T** and **R**. It relies on the induced fit model of binding, where a ligand binding induces structural changes in the enzyme. As an allosteric activator binds to the **T** state a series of sequential tertiary structural changes occur that can be transmitted between connected subunits, which results in a switch to the **R** state (Fersht, 1999) (Figure 1.13). The KNF model is also termed the sequential model due to this phenomenon. Unlike the MWC model structural changes conveyed by the KNF model can induce a negative structural change resulting in decreased binding of subsequent ligands (negative cooperativity), while the MWC model can only induce a population shift to the higher affinity **R** state (Fersht, 1999).

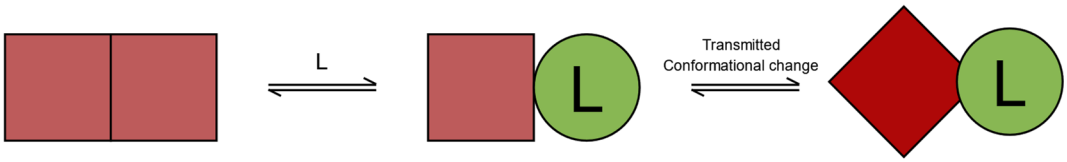

Figure 1.13. KNF Model. Allostery is mediated by ligand binding (L) inducing a sequential conformational change (from T (square) to R (circle)) that alters the enzymes' binding affinity for subsequent ligands. Adapted from Garrett and Grisham (2013).

1.3.1.3 Stereochemical Model

Both the MWC and KNF models did not offer any insight into the mechanisms of allosteric communication between the allosteric binding site and the active site. It was the advent of X-ray crystallography and structural biology that allowed new insights into the function of allostery. The first structure of haemoglobin (the classical model for allostery and the basis for the MWC and KNF models) was published by Max Perutz in 1960 (Perutz *et al.*, 1960). Following the later structures of *oxy*- and *deoxy*-haemoglobin Perutz hypothesised that allostery could be understood by careful structural analysis of allosteric enzymes (Motlagh *et al.*, 2014). Perutz showed that a small number of salt bridges mediate the conformational changes between allosteric states in haemoglobin (Hilser *et al.*, 2012). In addition to this it has also been hypothesised that allostery is mediated by conserved allosteric pathways that transmit communications between the allosteric and active sites (Motlagh *et al.*, 2014). It has been proposed that these pathways consist of a set residues that are in dynamic contact that transmit a perturbation at the allosteric site to the active site (del Sol *et al.*, 2009) (Figure 1.14).

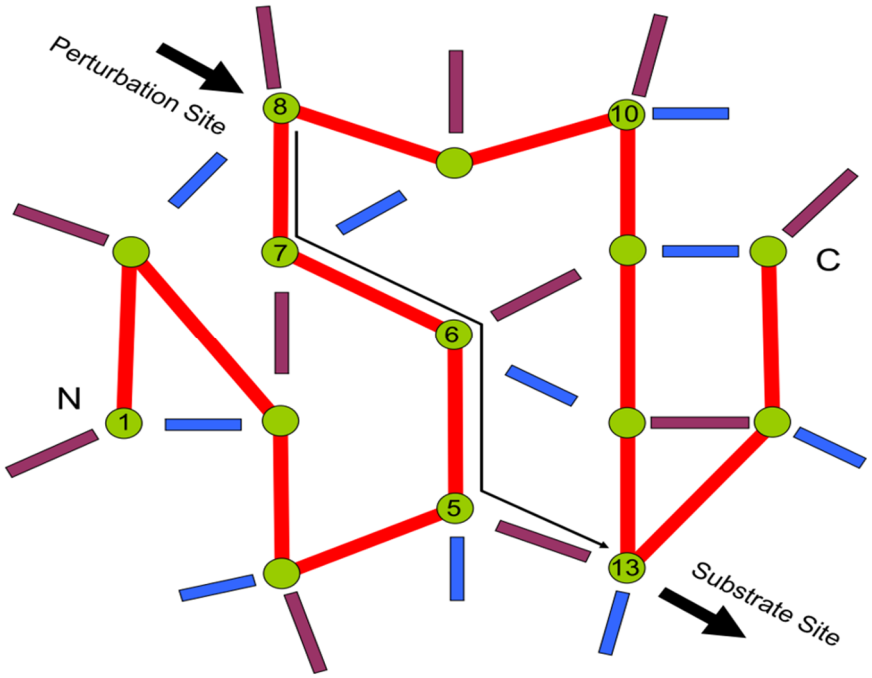

Figure 1.14. Allosteric pathways transmit perturbation at the allosteric site to the active site through conserved pathways (del Sol et al., 2009).

1.3.1.4 Limitations of structural approaches to allostery

Although these models have found widespread application, they do not represent a complete view of allostery, and do not offer any insights into allosteric mechanisms. Under these models allostery is mediated by structural shifts in oligomeric proteins and so is limited to oligomers, however allostery has been seen in monomeric proteins. The MWC model is limited by its requirement for symmetry. Certain allosteric regulators can be both activators or repressors under different conditions, a property not accounted for by current conformational models (Hilser *et al.*, 2012).

1.3.2 Protein dynamics approach to allostery

The role of dynamics as a mechanism for allostery rather than conformational changes was first hypothesised by Cooper and Dryden in 1984 (Cooper & Dryden, 1984). Enzymes were beginning to be seen as dynamic molecules that could undergo various forms of intramolecular motion, not simply the static molecules

that had previously been believed (Cooper & Dryden, 1984). Cooper and Dryden showed that allostery could arise from changes in conformational dynamics (statistical distribution of possible conformations) caused by changes in thermal fluctuations caused by ligand binding (Motlagh *et al.*, 2014). It was also noted by Cooper and Dryden that rate processes that utilise thermal fluctuations such as catalysis, and ligand binding and dissociation may be regulated by dynamic processes (Cooper & Dryden, 1984).

A paper published by Colombo and colleagues in 1992 showed the importance of dynamic effects in allostery (Colombo *et al.*, 1992). They showed that the structural transition involved in haemoglobin allostery also involved the additional binding of approximately 60 water molecules (Colombo *et al.*, 1992). This extra hydration stabilised the conformational change revealing the role of the entire molecule in the allosteric mechanism (Motlagh *et al.*, 2014).

1.3.2.1 Ensemble model

The ensemble allosteric model is a model which redistributes the two states found in the MWC model as a statistical ensemble (distribution of possible conformations). In the ensemble model an enzyme is treated as two separate domains that can interact (Hilser *et al.*, 2012). As a general case one domain is the allosteric domain and the other the catalytic domain. Each domain can independently sample multiple conformations based around a statistical average. In the simplest case there are two states the low affinity **T** state and the high affinity **R** state. When a conformational transition occurs in one domain, for example by binding a ligand, the probability of the other domain being in the low affinity **T** state is altered (Hilser *et al.*, 2012).

1.3.2.2 Molecular dynamics in allostery

Enzymes exist as an ensemble of different conformations (Marino *et al.*, 2018). Crystallography techniques present only a static view of protein conformation, so therefore do not represent a full view of enzyme dynamics. Molecular dynamics (MD) simulations can probe conformational variability at atomic level resolution (Marino *et al.*, 2018). Molecular dynamics uses Newtonian mechanics to predict

the position of atoms as a function of time rather than computationally intensive quantum mechanical models (Hertig *et al.*, 2016). Perturbations can then be introduced to the system and their effects studied (Hertig *et al.*, 2016). MD simulations can be used to study and predict the physical properties (such as stability and dynamics) of proteins or their interactions with other molecules.

1.4 The temperature dependence of reaction rates

The relationship between reaction rates and temperature was first formalised by Svante Arrhenius in the 19th century, based on empirical observations (Equation 1.4).

Equation 1.4. Arrhenius Equation

$$k = Ae^{-E_A/RT}$$

Arrhenius expanded on the earlier work of Jacoubus Henricus van't Hoff to describe the rates of simple chemical reactions including defining a term for the activation energy of a system (E_A) (Justi & Gilbert, 1999). Arrhenius described the rate as being exponentially related to the activation energy divided by temperature. A pre-exponential term was included to describe the frequency of collisions between reactant molecules. Arrhenius proposed that the necessary energy for a reaction to proceed was associated with an energy barrier, and it was the proportion of reacting molecules in this higher energy or 'active' state which conveyed the temperature dependence of chemical rates (Justi & Gilbert, 1999). Due to the exponential nature of the equation small changes in temperature result in an exponential change in reaction rate.

The Arrhenius equation was expanded independently by Michael Polanyi and Meredith Evans, and Henry Eyring in 1935. Aspects of previous thermodynamic, kinetic and statistical models were unified into a new comprehensive model, Transition State Theory (TST) (Justi & Gilbert, 1999). Under TST consideration was

given to the transition of molecules from reactant to product. Reactants were considered to pass through a continuous set of defined configurations between reactants and products (Evans, 1938). The rate of a chemical reaction was therefore redefined such that consideration was given to both the thermodynamic considerations of reaction rates though an intermediate transition state species, as well as the kinetics of molecule collisions (Justi & Gilbert, 1999). TST led to the development of the Eyring-Polanyi equation, which included additional terms, the Boltzmann constant (k_B), Planck's constant (h), and the Gibbs free energy change between the reactant state and transition state ($\Delta G^{\ddagger}$) (Equation 1.5).

Equation 1.5. Eyring equation

$$k = \left(\frac{k_B T}{h}\right) e^{-\Delta G^{\ddagger}/RT}$$

It is often assumed that $\Delta G^{\ddagger}$ is linearly dependant on temperature and thus $\Delta G^{\ddagger}$ varies with temperature according to the equation $\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\,\Delta S^{\ddagger}$ (Arcus *et al.*, 2016). The Eyring equation (Equation 1.5) can thus be represented as Equation 1.6 by the substitution of $\Delta G^{\ddagger}$ for its components, which is also equivalently represented as Equation 1.7.

Equation 1.6. Expanded Eyring equation

$$k = \left(\frac{k_B T}{h}\right)\left(e^{\frac{-\Delta H^{\ddagger}}{RT}}\right)\left(e^{\frac{\Delta S^{\ddagger}}{R}}\right)$$

Equation 1.7. Expanded Eyring equation

$$ln(k) = ln\left(\frac{k_B T}{h}\right) - \frac{\Delta H^{\ddagger}}{RT} + \frac{\Delta S^{\ddagger}}{R}$$

1.4.1 Temperature dependence of biological macromolecules

The Arrhenius and Eyring equations have traditionally been applied to biological systems as well as simple chemical systems. It has been widely accepted that the deviation from these models at elevated temperatures in biological systems is caused by irreversible thermal denaturation. An enzyme following the Arrhenius/Eyring models is expected to show an exponential increase in rate with temperature (Figure 1.15A), however it has been noted that many enzymes show much slower and non-exponential rate increases (Thomas & Scopes, 1998) (Figure 1.15B). Enzymes have also been observed to show a maximum rate (T_{opt}) and subsequent decrease with temperature before thermal denaturation becomes significant (Thomas & Scopes, 1998). This is especially true for many psychrophilic enzymes which show decreases in rate at temperatures well below their thermal denaturation temperatures (Feller & Gerday, 2003). This suggests that enzyme-catalysed rates have a more complex temperature dependence than is described by the Eyring equation (Arcus *et al.*, 2016).

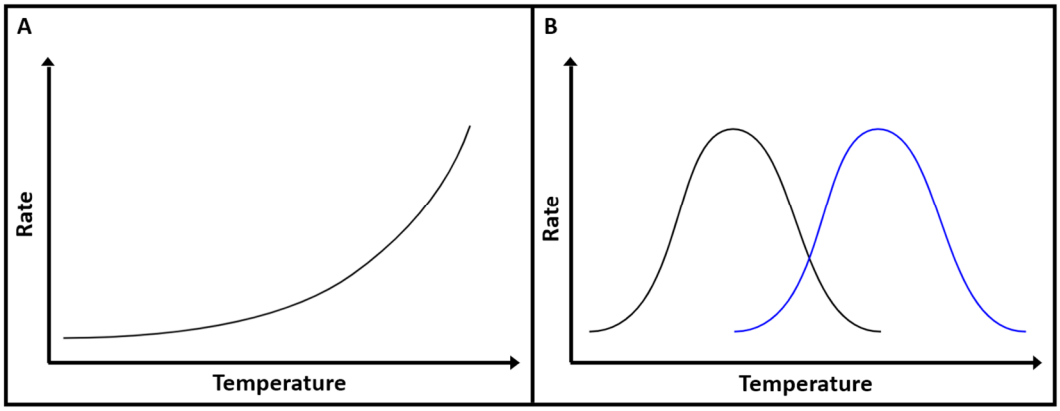

Figure 1.15. Models of Temperature dependence. A) Arrhenius and Eyring models produce an exponential increase in rate with temperature. B) Enzymes show a non-exponential rate increase with a maximal rate and subsequent decrease with temperature (psychrophilic (Black) and mesophilic (Blue) enzymes). Adapted from Arcus et al. (2016).

1.4.2 Macromolecular rate theory (MMRT)

An alternative explanation to model observed enzyme catalysed rates is MMRT, which incorporates a value for the activation heat capacity ($\Delta C_p^{\ddagger}$) into the rate equation (Hobbs *et al.*, 2013). Heat capacity can be defined as the ability of a system to absorb energy without changing temperature. Proteins are large molecules that have a correspondingly large heat capacity. It has been shown that the largest contribution to a protein's heat capacity is the number of accessible vibrational modes as well as a small contribution from the solvation interaction between the protein and the aqueous environment (Gómez *et al.*, 1995). Heat capacity defines the temperature dependence of the entropy and enthalpy of reaction based on Equation 1.8. Substituting the temperature dependence of entropy and enthalpy into the Eyring equation (Equation 1.7) gives the MMRT equation (Equation 1.9).

Equation 1.8. Temperature dependence of enthalpy and entropy

$$\Delta G_{T_0}^{\ddagger} = \left[\Delta H_{T_0}^{\ddagger} + \Delta C_p^{\ddagger}(T - T_0)\right] - T\left[\Delta S_{T_0}^{\ddagger} + \Delta C_p^{\ddagger}\left(\ln(T) - \ln(T_0)\right)\right]$$

Equation 1.9. MMRT equation

$$\ln(k) = \ln\left(\frac{k_B T}{h}\right) - \frac{\Delta H_{T_0}^{\ddagger} + \Delta C_p^{\ddagger}(T - T_0)}{RT} + \frac{\Delta S_{T_0}^{\ddagger} + \Delta C_p^{\ddagger}\left(\ln(T) - \ln(T_0)\right)}{R}$$

$\Delta C_p^{\ddagger}$ can also be temperature dependant. Adding a temperature dependant heat capacity gives Equation 1.10. This adds a new factor *A* which quantifies the temperature dependence of $\Delta C_p^{\ddagger}$.

Equation 1.10. MMRT equation with temperature dependent $\Delta C_P^{\ddagger}$

$$\ln(k) = \ln\left(\frac{k_B T}{h}\right) - \frac{\Delta H_{T_0}^{\ddagger} + \left[\Delta C_{p,T_0}^{\ddagger} + A(T - T_0)\right](T - T_0)}{RT} + \frac{\Delta S_{T_0}^{\ddagger} + \left[\Delta C_{p,T_0}^{\ddagger} + A(T - T_0)\right]\ln\left(\frac{T}{T_0}\right)}{R}$$

Heat capacity adds a new dimension of understanding to enzyme catalysed rates. $\Delta C_p^{\ddagger}$ is the difference in heat capacity between the enzyme-substrate complex and enzyme-transition state complex. This is influenced mostly by the change in the frequencies of vibrational modes between the two aforementioned states (Arcus *et al.*, 2016). Enzymes generally have a large negative $\Delta C_p^{\ddagger}$ due to tight binding of the enzyme to the transition state resulting in a shift of the vibrational modes to higher frequencies. MMRT produces a curved temperature dependence (Figure 1.16). The degree of curvature is defined by the value of $\Delta C_p^{\ddagger}$. As $\Delta C_p^{\ddagger}$ gets more negative the curvature of the temperature dependence becomes more prominent. At $\Delta C_p^{\ddagger} = 0$ the equation collapses back to an Arrhenius function (Hobbs *et al.*, 2013).

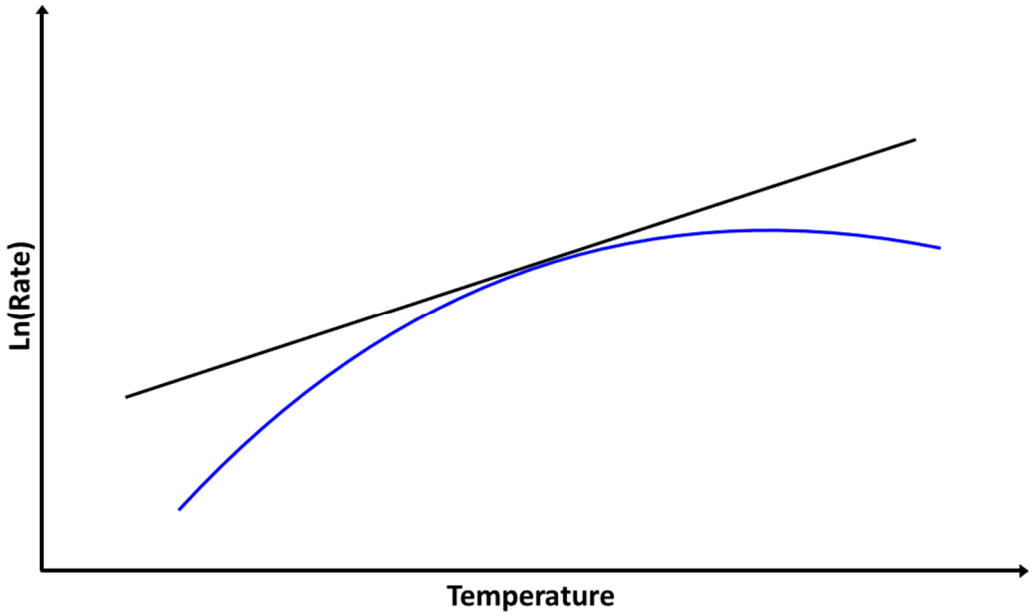

Figure 1.16. Temperature dependence of enzyme rates. Arrhenius and Eyring models (Black) produce a linear temperature dependence. MMRT (Blue) produces a curved temperature dependence. Degree of curvature is dependent on the value of $\Delta C_P^{\ddagger}$. Adapted from Arcus et al. (2016).

1.5 X-ray Crystallography

X-ray crystallography is a tool for probing the structure of molecules, including proteins. X-ray crystallography is based on the interaction of X-rays with elementary particles and utilises the scattering of X-rays by electrons (Drenth, 1999). The intensity of X-ray scattering is dependent on the number of electrons per atom (Blow, 2002). The X-rays used in X-ray diffraction (XRD) are monochromatic. When an X-ray hits an electron some of the energy is absorbed and reemitted at the same wavelength as a spherical wave. When this process occurs in a regular lattice such as a crystal, diffracted X-rays add up at specific angles relative to the angle of incident radiation (Blow, 2002). When distances in the lattice correspond to integer multiples of the incident wavelength diffracted X-rays will add up constructively. When diffracted X-rays line up in-phase of the waves they add up constructively. If they line up out of phase of the waves, they add up destructively (Figure 1.17) (Drenth, 1999). This is defined mathematically by Braggs law (Equation 1.11).

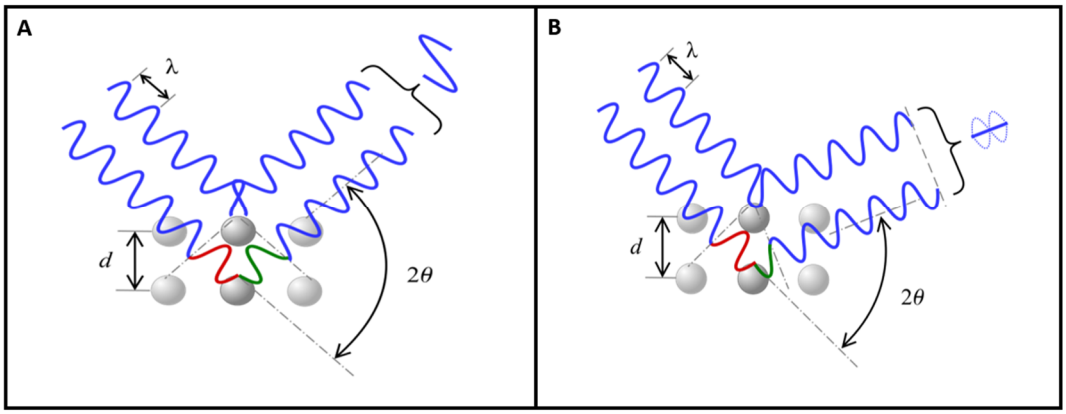

Figure 1.17. Braggs Law. A) Diffracted X-rays line up when the waves are in phase and add constructively, leading to amplification of signal. B) Diffracted X-rays line up out of phase of the waves and add up destructively. Adapted from Dang Ngoc Chan (2011).

Equation 1.11. Braggs Law

$$n\lambda = 2d \sin\theta$$

When crystals are irradiated with X-rays a diffraction pattern is produced (Blow, 2002). Each spot, called a reflection, corresponds to a particular set of planes of the crystal in reciprocal space. The position and intensity of each reflection gives information about the electron density in a given plane. By rotating the crystal, a full 3D picture of the electron density can be built (Blow, 2002). Electron density is calculated by performing a mathematical operation called a Fourier transform on the intensity data. The intensity is a product of wave interactions and is determined by the amplitude and phase of the diffracted radiation. The phase cannot be determined experimentally so must be determined by another method, such as molecular replacement, anomalous scattering or isomorphous replacement (Drenth, 1999).

1.5.1 Protein Crystallography

Due to the periodicity of a crystal the observed intensities are much stronger. The amplitude of diffracted X-rays is linearly dependant on the number of points in the scattering lattice (Blow, 2002). A well-formed, large crystal produces the best formed lattice to produce maximal diffraction. Therefore, significant effort is put into obtaining sufficient quality crystals for X-ray diffraction.

Protein crystals are grown from supersaturated conditions. The conditions should provide supersaturation for the protein but not significantly perturb its normal state (McPherson & Gavira, 2013). Protein crystals are usually grown in aqueous solution with additives and salts to promote crystallisation. A common method to achieve supersaturation is the ‘hanging-drop’ where a small drop containing protein and crystallisation solution is suspended above a reservoir solution containing only crystallisation solution (Figure 1.18). This sealed environment allows vapour diffusion to occur which results in a net movement of water into the reservoir leading to an increase in protein saturation.

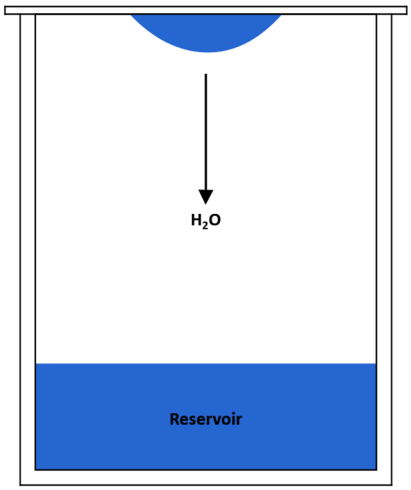

Figure 1.18. Hanging-drop vapour diffusion method of protein crystallisation. A mixed protein drop is suspended above an unadulterated reservoir solution. Vapour diffusion results in water loss increasing protein saturation and promoting crystallisation.

There are two stages to crystallisation, nucleation and growth. The nucleation stage is poorly understood. During this period small aggregates form which become the nuclei for crystal growth (McPherson & Gavira, 2013). It has been suggested that crystals arise from critical nuclei that form in a distinct liquid protein phase that forms in saturating protein solutions (Wolde & Frenkel, 1997). Protein crystals then grow from the formed nuclei. The attachment of new molecules occurs at steps on the crystals surface, as it is more energetically favourable to bind at these steps. Steps are formed by defects in the crystal order or at nuclei that form randomly on the crystal surface (Drenth, 1999).

The conditions under which crystals will grow is difficult to predict and typically relies on screening a large number of conditions to identify potential conditions (McPherson & Gavira, 2013). These identified conditions can then be systematically altered to optimise crystal growth. The three general factors for crystallisation conditions are pH, ionic strength and precipitant. Changing the pH alters the ionisation state of surface amino-acid residues and can alter the degree of attraction between protein molecules (McPherson & Gavira, 2013). Precipitants interact with the solvent in such a way that that the protein saturation is altered to promote crystallisation. Similarly, salts compete with proteins for interactions with water, promoting crystallisation (McPherson & Gavira, 2013).

1.6 Glucosidase MalL from *Bacillus subtilis* as a model enzyme

MalL is an enzyme involved in maltose utilisation in *Bacillus subtilis*. MalL was identified as a putative α-glucosidase encoded by the *yvdL* open reading frame, which was identified as part of the *B. subtilis* genome sequencing project (Schönert *et al.*, 1998). MalL was characterised as a specific Oligo-1,4-1,6-α-glucosidase (sucrase-isomaltase-maltase). The activity of MalL can be followed spectroscopically at 405 nm by the cleavage of *p*-nitrophenyl-α-D-glucopyranoside (PNG).

MalL is part of the glycoside hydrolase superfamily of enzymes. Based on structural and mechanistic similarities glycoside hydrolases can be further divided. MalL and other Oligo-1,6-α-glucosidases are members of the group 13 glycoside hydrolases (GH-13) or the GH-H clan. GH-13 enzymes share a conserved structure. The catalytic core (domain A) consists of an 8-stranded β/α barrel first described in the enzyme triosephosphate isomerase (TIM). A second domain (B) protrudes between β_3 and α_3 and forms part of the active site structure. GH-13 enzymes also contain a third (C) Greek-key all β carboxy-terminal domain (Figure 1.19).

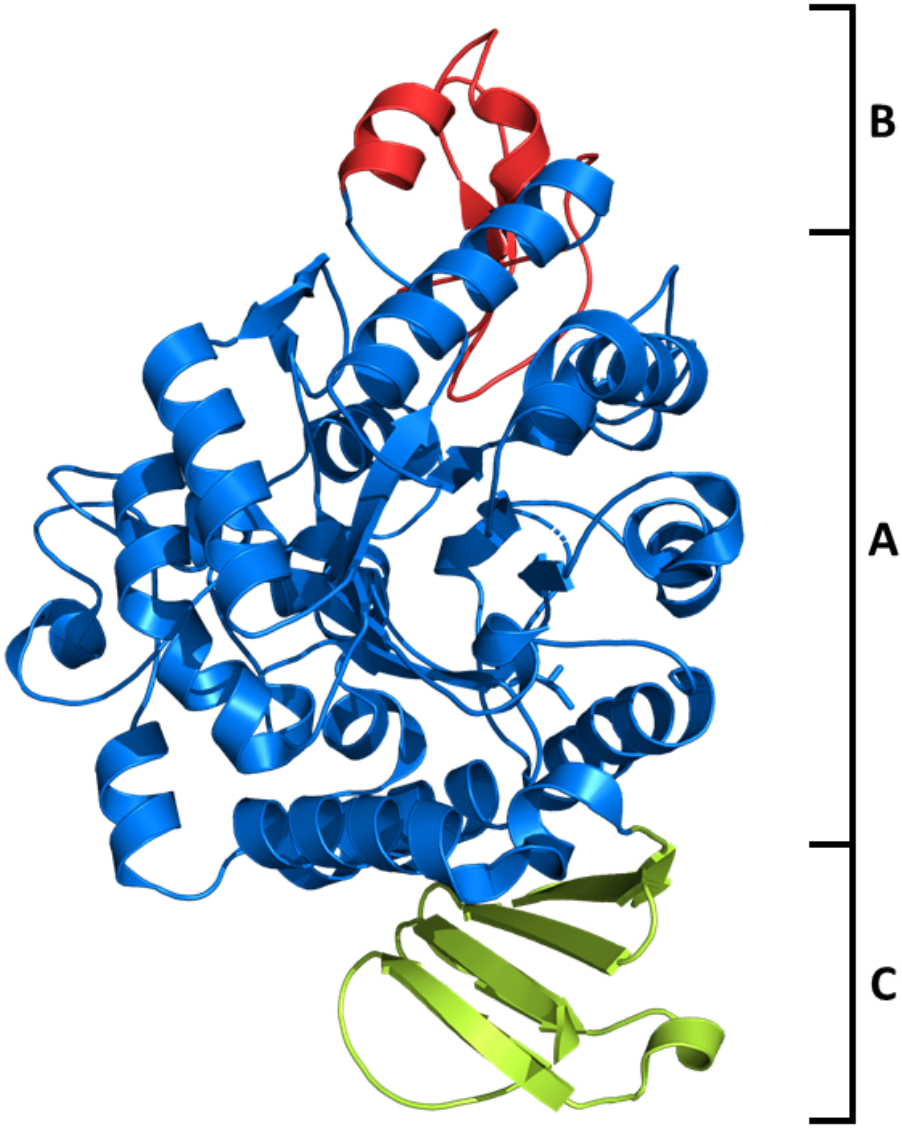

Figure 1.19. Domain architecture of the Bacillus subtilis isomaltase enzyme MalL. Domain A (Blue) is the catalytic TIM barrel. Domain B (Red) is a β₃ to α₃ extension. Domain C (Green) is a C-terminal all-β Greek key domain.

MalL has been extensively used as a model for MMRT. MalL is a relatively large monomeric enzyme which lends itself well to studying dynamic effects (such as during MMRT). MalL is particularly stable, and its activity can be easily measured using a direct assay. Various MalL mutants have been well characterised during previous investigations which have contributed to the development of the MMRT enzymatic model (Figure 1.20).

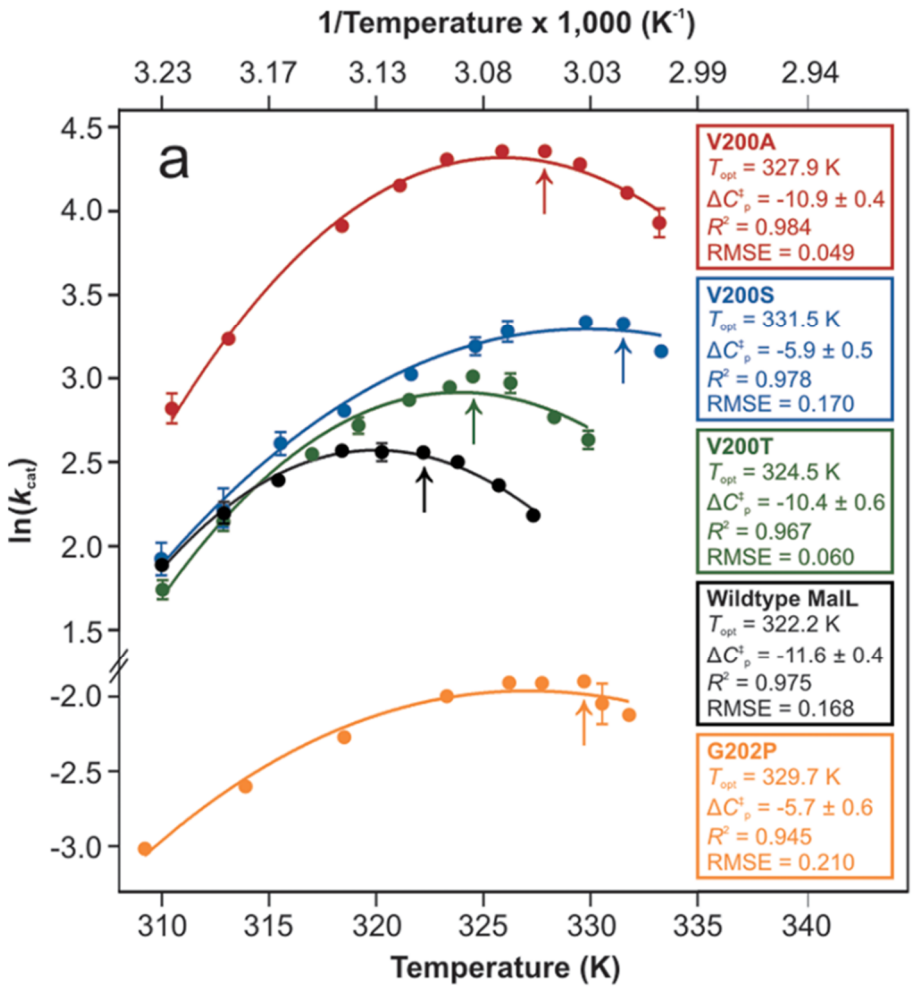

Figure 1.20. Characterisation of a range of individual MalL mutants have contributed to the development of the MMRT enzymatic model (Hobbs et al., 2013). Arrows indicate the T_{opt} for each enzyme.

In addition to this characterisation, 3D structures (achieved via protein crystallography) have been obtained for wild type (WT) MalL, the V200A, V200S and G202P variants as well as MalL bound to the transition state analogue 1-deoxynojirimycin (Hobbs *et al.*, 2013; van der Kamp *et al.*, 2018) and bound to urea (unpublished). Acquiring a range of 3D structures of MalL thus far has contributed to the knowledge of the dynamic mechanisms of this enzyme and to the development of MMRT.

1.7 Aims of MSc project

The aim of this project was to use MMRT as a tool to determine how allosteric enzyme regulation may be interpreted by changes in enzyme heat capacity along the reaction coordinate. To test this experimentally we use the model enzyme MalL and generate and characterise of a range of MalL mutants. Mutants are characterised by using activity assays to determine Michaelis-Menten kinetics and over a range of temperatures in order to be modelled using MMRT. This allows us to determine heat capacity changes and infer changes to the enzyme's dynamics along the reaction coordinate. In addition, protein crystallography techniques are used to obtain detailed 3D structural information for the MalL mutants.

2 Materials and methods

2.1 Protein production and purification

2.1.1 Buffers and Media

Composition of general buffers and media used in all experiments is given in Table 2.1.

Table 2.1. General Buffers and Media

Solution	Components
LB media	10 g L^{-1} NaCl, 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract
LB Agar	10 g L^{-1} NaCl, 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 15 g L^{-1} agar
Lysis buffer	10 mM KCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 25 mM imidazole, pH 7.0
Elution buffer	10 mM KCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.5 M imidazole, pH 7.0
Size exclusion buffer	20 mM HEPES, pH 7.8
Assay buffer	40 mM Na_2PO_4, 150 mM NaCl, pH 8.0

2.1.2 Protein Expression

A glycerol stock of DH5α *Escherichia coli* containing the *Bacillus subtilis* isomaltase gene (MalL) (wild type and S536R variant) in the pPROEX-Htb expression vector was provided by E. Prentice. Single colonies were isolated by plating on a LB agar plate with 0.1 mg mL^{-1} ampicillin and grown overnight at 37°C.

2.1.2.1 Glycerol Stocks

Glycerol stocks were made adding equal volumes of 50% v/v glycerol and bacterial culture and stored at -80°C.

2.1.2.2 Expression protocol

Glycerol stocks were used to inoculate 20 mL starter cultures in LB media with 0.1 mg mL^{-1} ampicillin. Cultures were grown overnight at 37°C with shaking at 180 rpm. Starter cultures were used to seed 1 L LB expression cultures in a 2 L baffled flask with 0.1 mg mL^{-1} ampicillin and 0.2% w/v glucose. Expression cultures were grown at 37°C with shaking at 180 rpm until OD_{600} was approximately 0.5-0.7. Protein expression was induced with 0.75 mM IPTG and expressed at 18°C with shaking at 180 rpm overnight. Cells were harvested by centrifugation at 4542 g for 20 min at 4°C. Pelleted cells were stored at -20°C until purification.

2.1.3 Protein purification

Cells pellets were defrosted on ice and resuspended in approximately 20 mL Lysis buffer by vortexing. Cells were lysed on ice by sonication at level five with three 30 second bursts with 30 seconds of cooling between each sonication (Ultrasonic processor XL, Misonix, USA). Soluble protein was obtained in the supernatant by centrifugation at 13946 g for 20 minutes at 4°C. Protein containing supernatant was filtered through 1.2, 0.45, and 0.2 µm Minisart syringe filters (Sartorius, Germany) before being purified by immobilised metal affinity chromatography.

2.1.3.1 Immobilised metal affinity chromatography (IMAC)

Protein was purified using IMAC based on the N-terminal hexa-His tag affinity for Nickel. The HisTrap FF (5 mL, GE Healthcare, UK) column was prepared by stripping with two column volumes of EDTA (100 mM, pH 8.0) and then washing with two column volumes of MQ H_2O. The column was then recharged with one column volume of 100mM $NiCl_2$ and washed with two column volumes of MQ H_2O. The column was then equilibrated with Lysis buffer before loading filtered protein.

Protein elution was performed on an Äkta Basic or Purifier FPLC system. Non-target protein was washed from the column using 4% Elution buffer and 96% Lysis buffer at 1 mL min^{-1}, until absorbance at 280 nm had plateaued. Protein was

eluted from the column using a 0-100% gradient over 50 mL at 1 mL min^{-1} and collected in 2 mL fractions. Elution was followed by UV trace at 280 nm.

2.1.3.2 Size exclusion chromatography (SEC)

SEC was performed on a HiLoad 16/60 Superdex 200 pg (GE Healthcare, UK) column on a Äkta Basic FPLC. The column was equilibrated with size exclusion buffer. Protein was concentrated to approximately 5 mL prior to loading onto the column via a 5 mL injection loop. Elution was performed using size exclusion buffer at 0.5 mL min^{-1} and protein was collected in 2 mL fractions. Elution was followed by UV trace at 280 nm.

2.1.4 Polyacrylamide gel electrophoresis (PAGE)

All PAGE gels were made as native gels (no SDS) and run with an SDS running buffer. Native PAGE gels were made with 12% acrylamide resolving layer with a 5% acrylamide stacking layer. Gels were made five at a time according to the scheme in Table 2.2. Gels were run with 1 X Tris Glycine-SDS (Sodium dodecyl sulphate) running buffer. Samples were prepared with 1:4 of 4 X SDS loading dye. Buffer compositions are given in Table 2.3.

Table 2.2. Components of PAGE Gel

Component	Resolving gel (mL)	Stacking gel (mL)
MQ H_2O	10.35	8.625
30% w/v Acrylamide	12	2.125
Buffer (Resolving or Stacking)	7.5	1.6
10% w/v APS	0.15	0.0063
TEMED	0.0015	0.00063

Table 2.3. Buffer compositions for PAGE

Solution	Composition
4 X SDS loading dye	250 mM Tris (pH 6.8), 20% glycerol, 4% SDS, 10% mercaptoethanol, 0.025% w/v bromophenol blue
TG-SDS running buffer	25 mM Tris, 250 mM glycine, 0.1% SDS
Resolving buffer	1.5 M Tris, pH 8.0
Stacking buffer	1.0 M Tris, pH 6.8

Samples were stained by Fairbanks A staining solution (0.05% Coomassie blue, 25% isopropanol, 10% acetic acid) (Wong *et al.*, 2000). Background staining was removed using Fairbanks D stain (10% acetic acid). Gel images were captured using the Omega Lum G imaging system (Aplegen, USA) and processed using RawTherapee 5.7 software and annotated in Microsoft PowerPoint.

2.1.5 Measuring protein concentration

Protein concentration was measured by UV at 280 nm using the Nanodrop 2000 spectrophotometer (ThermoFischer Scientific, USA). Protein concentration was corrected with the extinction coefficient as calculated by ProtParam (Gasteiger *et al.*, 2005) based on the protein sequence.

2.1.6 Protein concentration

Protein was concentrated to the appropriate volume or concentration using Amicon Ultra (Merck Millipore, USA) spin concentrators with a 10 kDa molecular weight cut-off, by centrifugation at 2400 g.

2.2 Kinetic assay procedure

Enzymatic activity of MalL can be followed by the cleavage of *p*-nitrophenyl-α-D-glucopyranoside (PNG) into glucose and *p*-nitrophenol. The production of *p*-nitrophenol can be followed spectroscopically at 405 nm. Assays were performed with a 1:1 ratio of enzyme in size exclusion buffer and substrate (PNG) in assay buffer. Additional assay components (GuHCl, indole; Sigma-Aldrich, USA) were included at the relevant concentration with substrate in assay buffer. A molar extinction coefficient of 13270 M^{-1} (405 nm, pH 8) was used in subsequent kinetic calculations (Kraakman, 2017).

2.2.1 Michaelis-Menten Assay

Michaelis-Menten assays were performed on the stopped-flow spectrophotometer (Tgk Scientific, UK). MalL at 0.026 mg mL^{-1} in size exclusion buffer was mixed 1:1 with appropriate substrate concentration in assay buffer, resulting in a final enzyme concentration of 0.013 mg mL^{-1}. Temperature control was achieved using a circulating water bath connected to the observation cell on the stopped-flow spectrophotometer. Accurate temperature measurements were achieved using a thermocouple. Reaction was repeated in quadruplicate and where the first measurement deviated from subsequent measurements was discarded. Data was collected for 20 seconds in 0.039 second intervals.

2.2.2 Temperature Optimum Assays

Temperature assays were performed on the stopped-flow spectrophotometer (Tgk Scientific, UK). MalL at 0.026 mg mL^{-1} was mixed 1:1 with 2 mM PNG to give a final concentration of 0.013 mg mL^{-1} protein and 1 mM PNG (approximately 10 X K_M for all MalL variants). Temperature control was achieved with a circulating water bath connected to the stopped-flow observation cell. Accurate temperature measurements were obtained using a thermocouple. Reaction was repeated in

quadruplicate with a non-recorded dummy shot between each replicate. Data was collected for 20 seconds in 0.039 second intervals.

2.2.3 Data Analysis

Data was analysed by linear regression for the first 10 seconds of rate data using Tgk Scientific Kinetic Studio software. Data analysis and equation fitting was performed using GraphPad Prism 8.3.0.

2.2.3.1 Fitting of MMRT equation

The temperature **independent** $\Delta C_P^{\ddagger}$ MMRT equation was fit over a limited temperature range which included the highest enzyme rates. The reference temperature (T_0) was them set as the temperature with the maximum rate minus four degrees. Using variables from the temperature **independent** MMRT equation and the T_0, the temperature optimum was then calculated using Equation 2.1, based on this fit.

Equation 2.1. Calculation of T_{opt}

$$T_{opt} = \frac{\Delta \mathrm{H}_{T_0}^{\ddagger} - \Delta C_P^{\ddagger}.T_0}{\Delta C_P^{\ddagger} - R}$$

The temperature **dependant** MMRT equation was then fit to the full range of data with the reference temperature (T_0) set as T_{opt} minus 4 degrees.

2.3 Mutagenesis and cloning

MalL WT and S536R DNA were provided by E. Prentice. MalL TM and W531A/G were generated as described in Section 2.3.1.

2.3.1 Mutant specific cloning

2.3.1.1 MalL TM

The mutant MalL TM was synthesised using Invitrogen GeneArt Gene Synthesis (ThermoFischer Scientific, USA). MalL TM was synthesised in a pMA-T cloning vector. The MalL TM gene was isolated from the pMA-T cloning vector by restriction enzyme digest (Section 2.3.3) with PstI and NcoI and purified by agarose gel electrophoresis (Section 2.3.2). The pPROEX-Htb vector was similarly digested and purified. The purified MalL TM gene was ligated (Section 2.3.4) into the pPROEX-Htb vector and transformed (Section 2.3.5) into *E. coli* DH5α.

2.3.1.2 MalL W531A and W531G

MalL W531G and W531A were prepared by two-halves PCR mutagenesis, comprising of two PCR reactions. The first PCR reactions amplifies the gene in two sections introducing the desired mutation in the overlapping region. For each mutant a forward and reverse mutagenesis PCR was run. Mutation primers were ordered from Integrated DNA Technologies (IDT inc., USA). Reactions consisted of 4 µL of 5 X HOT FIREPol Master Mix (Solis BioDyne, Estonia), 0.5 µL of each primer (10 pmol μL^{-1}), 1 µL of MalL WT pPROEX-Htb template (72 ng µL), made up to 20 µL with ultrapure water.

Primers used are given in Table 2.4 and Table 2.5. Forward mutagenesis used forward pPROEX-Htb primers and reverse mutagenesis primers. Reverse mutagenesis used reverse pPROEX-Htb primers and forward mutagenesis primers. PCR cycling parameters are given in Table 2.6.

Table 2.4. pPROEX-Htb primers

Primer	Sequence (5' – 3')
pPROEX-Htb Fwd	AGC GGA TAA CAA TTT CAC ACA GG
pPROEX-Htb Rev	TAT CAG GCT GAA AAT CTT CTC

Table 2.5. Mutagenesis Primers for MalL W531G and W531A mutants. Bolded residues show the mutation location.

Primer	Sequence (5' – 3')
MalL W531G Fwd	CTGATTCATGAACGT**GGG**AAAGTTCTGATTAGC
MalL W531G Rev	GCTAATCAGAACTTT**CCC**ACGTTCATGAATCAG
MalL W531A Fwd	CTGATTCATGAACGT**GCG**AAAGTTCTGATTAGC
MalL W531A Rev	GCTAATCAGAACTTT**CGC**ACGTTCATGAATCAG

Table 2.6. PCR cycling parameters for MalL W531G and MalL W531A mutagenesis PCR reactions

Step	Temperature (°C)	Step Length (Fwd)	Step Length (Rev)
Pre-Denaturation	95	15 minutes	15 minutes
Denaturation	95	15 seconds	15 seconds
Annealing	60	45 seconds	45 seconds
Extension	72	2 minutes	15 seconds
Post-extension	72	10 minutes	10 minutes

The purity of each PCR was assessed by agarose gel electrophoresis (Section 2.3.2). A total of four PCR reactions for each mutagenesis reaction was pooled and purified by agarose gel electrophoresis (Section 2.3.2). Forward and reverse mutagenesis products were used as templates for a second PCR to produce a complete copy of the mutated gene using pPROEX-Htb primers. Reactions consisted of 4 µL of 5 X HOT FIREPol Master Mix (Solis BioDyne), 0.5 µL of each primer (10 pmol μL^{-1}), 0.3 µM of the forward and reverse mutagenesis PCR products, made up to 21 µL with ultrapure water. PCR cycling parameters are the same as given for the forward mutagenesis PCR given in Table 2.6 above.

The PCR products of the gene amplified in two sections were purified from solution using the High Pure PCR Product Purification Kit (Roche Diagnostics) and purity checked by agarose gel electrophoresis (Section 2.3.2). The subsequent joining PCR product was purified by agarose gel electrophoresis (Section 2.3.2). Purified products were used as template for an additional PCR to further amplify

products. Reactions consisted of 20 µL of 5 X HOT FIREPol Master Mix (Solis BioDyne), 2.5 µL of each primer (10 pmol μL^{-1}), 13.5 µL of W531G join PCR product (24.4 ng µL) or W531A join PCR product (32.6 ng μL^{-1}), made up to 100 µL with ultrapure water. PCR cycling parameters were the same as given for the forward mutagenesis PCR given in Table 2.6.

PCR products were purified from solution using the High Pure PCR Product Purification Kit (Roche Diagnostics) and purity checked by agarose gel electrophoresis (Section 2.3.2). The W531A and W531G genes were digested (Section 2.3.3) with PstI and NcoI and purified by agarose gel electrophoresis (Section 2.3.2). The pPROEX-Htb vector was similarly digested and purified. The purified MalL W531G and W531 A genes were ligated (Section 2.3.4) into the pPROEX-Htb vector and transformed (Section 2.3.5) into *E. coli* DH5α.

2.3.2 Agarose gel electrophoresis

Agarose gels were made up as required to between 0.8 – 2.0% w/v agarose in TAE buffer (40 mM Tris-acetate, 1 mM EDTA). Invitrogen SYBER safe at a final 1 X concentration (ThermoFischer Scientific, USA; 10000 X) was added to allow for DNA visualisation. Samples were loaded with 5 X DNA loading dye [25% v/v glycerol, 0.2% w/v bromophenol blue]. Samples were visualised under blue light. Sizes were analysed by comparison to the Invitrogen 1 Kb Plus DNA ladder (ThermoFischer Scientific, USA). Gel images were collected using the Omega Lum G gel imaging system (Aplegen, USA) and processed using RawTherapee 5.7 and Microsoft PowerPoint.

2.3.2.1 Purification from agarose gel

Relevant bands were excised from the agarose gel by visualisation under blue light. DNA was purified from the agarose gel using either the QIAquick Gel Extraction Kit (QIAGEN, Germany) or the UltraClean 15 DNA Purification Kit (MO BIO Laboratories, USA) according to the manufacturer's instructions. DNA was eluted into between 15 and 50 µL of elution buffer.

2.3.3 Restriction Enzyme Digests

Restriction enzyme digests were performed using the NEB Cutsmart system (New England Biolabs, USA). All digests performed utilised NEB PstI-HF and NcoI-HF (20000 U mL^{-1}) restriction enzymes. Digests were performed with 2000-3000 ng DNA, 60 U of each restriction enzyme and the appropriate amount of 10 X Cutsmart buffer in a final reaction volume of 50 µL. Digest reactions were left at 37°C for 2 hours. Non-target DNA from the restriction digests was separated by agarose gel electrophoresis (Section 2.3.2) and target DNA was purified from the gel (Section 2.3.2.1).

2.3.4 Ligations

Ligation reactions were performed using NEB T4 DNA ligase (400000 U mL^{-1}) (New England Biolabs, USA). Ligation reactions were done with a molar insert to vector ratio of between 2:1 and 3:1. Reactions were left at 4°C for 20 minutes and 16°C for 2 hours.

2.3.5 Transformation of *E. coli* by heat shock

Plasmids were transformed into chemically competent *E. coli* DH5α cells by heat shock. Chemically competent cells were defrosted on ice. The ligation reaction (10 µL) was added to chemically competent cells (100 µL) and incubated on ice for 30 minutes. Cells were heat shocked at 42°C for 45 seconds and then incubated on ice for a further 2 minutes. One millilitre of SOC media [20 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 585 mg L^{-1} NaCl, 186 mg L^{-1} KCl, 2.46 g L^{-1} $MgSO_4.7H_2O$, 2.03 g L^{-1} $MgCl_2.6H_2O$, 20 mM glucose (sterile filtered)] was added and the cells were incubated at 37°C for 60 minutes. Aliquots of transformed cells were plated on LB Agar plates with 0.1 mg mL^{-1} ampicillin. Colonies were screened for successful mutation by colony PCR (Section 2.3.6).

2.3.6 Colony PCR

Transformation colonies were screened for successful insert by PCR using pPROEX-Htb primers (Table 2.4) flanking the gene insert position. Single colonies were picked from a plate using a sterile pipette tip, streaked onto an LB agar plate containing 0.1 mg mL^{-1} ampicillin for preservation and then the cell remnants were resuspended in the PCR reaction mixture for use as template DNA. PCR reactions were set up for each colony using either HOT FIREPol (SolisBioDyne, Estonia) or Invitrogen Taq Polymerase (ThermoFischer Scientific, USA). PCR results were assessed by agarose gel electrophoresis (Section 2.3.2). Reactions consisted of 4 µL of 5 X HOT FIREPol Master Mix (Solis BioDyne, Estonia), 0.5 µL of each primer (10 pmol μL^{-1}), made up to 20 µL with ultrapure water for HOT FIREPol. Composition of Taq PCR reactions is given in Table 2.7. PCR cycling conditions are given in Table 2.8.

Table 2.7. Colony PCR composition for Taq polymerase

Component	Volume	Final Concentration
10 X PCR buffer	2.5 µL	1 X
50 mM $MgCl_2$	0.5 µL	1 mM
10 mM dNTPs	0.5 µL	0.2 mM
pPROEX-Htb Fwd primer (10 pmol μL^{-1})	0.75 µL	0.3 µM
pPROEX-Htb Rev primer (10 pmol μL^{-1})	0.75 µL	0.3 µM
Taq polymerase (5 U μL^{-1})	0.25 µL	1.25 U
Ultrapure Water	to 25 µL	-

Table 2.8. PCR cycling parameters for colony PCR reactions

	HOT FIREPol		Taq polymerase	
Step	**Temperature**	**Step Length**	**Temperature**	**Step Length**
Pre-Denaturation	95	15 minutes	94	2:45 minutes
Denaturation	95	15 seconds	94	15 seconds
Annealing	60	45 seconds	58	30 seconds
Extension	72	2 minutes	72	2 minutes
Post-extension	72	10 minutes	72	7 minutes

The LB agar plate with colonies selected for colony PCR was incubated overnight at 37°C. Colonies with a successful gene insert were picked from the fresh plate and grown overnight in 5.5 mL LB media with 0.1 mg mL^{-1} ampicillin. Plasmid DNA was isolated from 5 mL of this culture (Section 2.3.7.2) and sent for sequencing to confirm mutation and gene integrity (Section 2.3.8). The remaining 0.5 mL was used to create a glycerol stock (Section 2.1.2.1).

2.3.7 DNA purification

2.3.7.1 Purification from PCR

Purification of DNA from PCR solution was achieved using the High Pure PCR Product Purification Kit (Roche Diagnostics, USA) according to the manufacturer's instructions. Where significant non-target products were observed DNA was additionally purified on by separation on agarose gel (Section 2.3.2).

2.3.7.2 Plasmid purification

E. coli DH5α (from glycerol stock or transformation plate) containing the relevant plasmid was grown overnight in LB media (5 mL) containing 0.1 mg mL^{-1} ampicillin. Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (QIAGEN, Germany) according to the manufacturer's instructions. Plasmid DNA was eluted into 30 µL elution buffer and stored at -20°C.

2.3.8 Sequencing

Plasmid DNA was extracted (Section 2.3.7.2) and sent to the Massey Genome Sequencing Service with samples prepared with pPROEX-Htb forward and reverse primers (Table 2.4)] according to the sample guidelines.

2.4 Crystallisation

Protein was expressed and purified as described in section 2.1 and concentrated to between 8 and 12 mg mL^{-1}. MalL S536R was concentrated to 10 mg mL^{-1}. MalL triple mutant (TM) was concentrated to between 8 and 12 mg mL^{-1}. The final successful crystallisation concentration for MalL TM was 12 mg mL^{-1}.

2.4.1 Initial crystallisation condition screening

Initial crystallisation conditions were determined using the sitting drop method. Pre-prepared crystal screens (Crystal Screen HT-HR2-130, Index-HR2-144, PEGRx HT-HR2-130, SaltRx-HR2-136) from Hampton Research were used as precipitant solutions (The salt screen was omitted from the MalL TM crystallisation screen due to limited protein). Aliquots of 100 µL of precipitant solution were pipetted into 96-well Intelli-Plate 96-2 Low Profile crystallisation plates (Hampton research, USA). The Mosquito crystallisation robot (TTP LabTech Ltd, UK) was used to dispense 200 nL crystallisation drops consisting 1:1 of protein to precipitant solution. Trays were sealed with Crystal Clear Sealing Film (Hampton Research, USA) and left at 18°C.

2.4.2 Fine screening of crystallisation conditions

Initial conditions identified were further refined by using a hanging-drop vapour diffusion technique. Conditions were varied from the initial conditions by altering pH, precipitant concentration, ionic strength and hanging-drop size. Drop sizes of between 1 and 4 µL with varying ratios of protein to precipitant solution were

suspended above 500 µL of precipitant solution on a 24 well VDX plate (Hampton Research). Wells were sealed with Glisseal N grease (Borer chemie, Switzerland) and left at 18°C.

2.4.3 Final crystallisation conditions

2.4.3.1 MalL S536R

Final crystallisation conditions for MalL S536R were 0.1 M Tris pH 8.0, 0.2 M ammonium acetate and 18% w/v PEG 10,000. Protein concentration was 10 mg mL^{-1} and drop size was 4 µL with a 1:1 ratio of protein to precipitant solution.

2.4.3.2 MalL triple mutant (TM)

Final crystallisation conditions for MalL TM were 0.1 M Tris HCl pH 8.5, 0.22 M Sodium Acetate, 27% w/v PEG 4,000. Protein concentration was 12 mg mL^{-1} and drop size was 3 µL consisting of 2 µL precipitant solution and 1 µL protein.

2.4.4 Data Collection

2.4.4.1 Crystal preparation

Crystals were removed from crystallisation drops with a cryo-loop (Hampton Research, USA) and transferred to a cryoprotectant solution. Cryoprotectant for MalL S536R was 0.1M Tris pH 8.0, 0.2 M ammonium acetate and 17% w/v PEG 10,000 with 20% v/v glycerol. Cryoprotectant for MalL TM was the same composition as the crystallisation precipitant solution with 20% v/v glycerol. Crystals were then frozen in liquid nitrogen for data collection.

2.4.4.2 Data collection

X-ray diffraction data was collected on the MX2 beamline at the Australian Synchrotron, Melbourne with an EIGER X 16M pixel detector (Dectris Ltd, Switzerland)

2.4.5 Data Processing

2.4.5.1 Indexing and integration

Data was indexed, integrated and scaled using XDS (Kabsch, 2010). Unit cell parameters and space group were determined using Pointless (Evans, 2011).

2.4.5.2 Scaling

Data was scaled and merged in AIMLESS (Evans & Murshudov, 2013) within the CCP4 software suite (Winn *et al.*, 2011). Data was examined and rescaled to optimise I/σI to approximately 2 in the outer resolution shell.

2.4.5.3 Matthews coefficient

The number of molecules in the asymmetric unit was calculated using the Matthews coefficient (Kantardjieff & Rupp, 2003; Matthews, 1968) within the CCP4 software suite (Winn *et al.*, 2011).

2.4.5.4 Molecular replacement

Molecular replacement was performed using the previously solved apo MalL WT structure (PDB: 4M56) for both the MalL S536R and MalL TM datasets. Molecular replacement for MalL S536R was performed using Molrep (Vagin & Teplyakov, 1997) in the CCP4 software suite (Winn *et al.*, 2011). Molecular replacement for MalL TM was performed using PHASER (McCoy *et al.*, 2007) in the PHENIX software suite (Adams *et al.*, 2010).

2.4.6 Model refinement

The structure was refined manually in COOT (Emsley *et al.*, 2010) with the $2F_O$-F_C electron density map contoured to 1 σ. Model refinement was completed in phenix.refine (Afonine *et al.*, 2012) in the PHENIX software suite or Refmac5 (Murshudov *et al.*, 1997) in the CCP4 software suite (Winn *et al.*, 2011).

2.4.7 Structure analysis and visualisation

B-factors were calculated using Baverage (Dodson, 1991) within CCP4. Root Mean Square Deviation (RMSD) values were calculated using porSMART (Nicholls *et al.*, 2014). Pymol (Schrödinger, 2000) was used for visualisation of structures and to generate images.

2.5 Normal mode analysis

Normal mode analysis to predict protein stability changes was performed using the Dynamut web server using the C-alpha force field (Rodrigues *et al.*, 2018).

2.6 Figures

Figures were made using Gravit Designer (Corel Corporation, 2019), Microsoft PowerPoint (Microsoft, 1990) and ChemDraw (Perkin Elmer Inc., 1998).

3 Mutation of MalL to mimic urea binding and demonstrate allostery

3.1 Introduction

Although allostery has been described in the literature for many years, the precise physical mechanisms by which allostery acts is not fully understood. The research presented here aims to apply macromolecular rate theory (MMRT) to describe a potential mechanism for allosteric action.

MMRT describes the response of enzymatic rate to temperature and adds the change in heat capacity ($\Delta C_P^{\ddagger}$) as a variable of enzymatic rate (Hobbs *et al.*, 2013). An enzymes' heat capacity is mainly described by the number and energy of vibrational modes of the molecule. Using this relationship, we can say that changes in heat capacity will directly relate to changes in energies of the vibrational modes of the enzyme, which in turn may relate to the enzyme rate. It has been well noted that enzymes often become more rigid in their ligand bound states accompanied by a shift in the vibrational modes to higher frequencies and therefore a change in heat capacity, termed ΔC_P (Arcus *et al.*, 2016). This therefore presents a potential mechanism by which ligand binding can alter enzymatic rate through changes in $\Delta C_P^{\ddagger}$. We thus sought to explore the role of $\Delta C_P^{\ddagger}$ changes in allosteric regulation.

Allosteric regulation was explored by characterising the kinetics of the enzyme isomaltase (MalL). Previous work (unpublished) has characterised the effect of non-denaturing (or low) concentrations of urea on the rate of MalL across a range of temperatures (Figure 3.1).

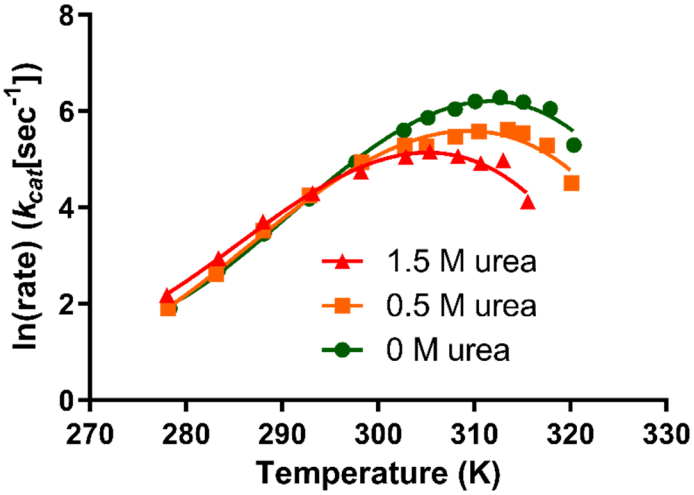

Figure 3.1. The temperature optima for MalL is decreased with increasing amounts of urea added to the kinetic assay (unpublished). The solid line represents the fit of the MMRT equation.

An obvious outcome of these experiments was that the temperature optima for MalL decreased in increasing amount of urea. Varying the concentration of urea also altered the dynamics of MalL activity across the temperature range, which was apparent after fitting the MMRT equation to the data. Interestingly at 1.5 M urea concentration, urea acted as a rate activator at low temperatures and a rate inhibitor at high temperatures. To probe the structural effects of urea binding on enzyme rate, MalL was crystallised, and the structure solved with varying concentrations of urea (unpublished). Interestingly, these structure results identified a urea binding pocket in the C-terminal domain of MalL (Figure 3.2).

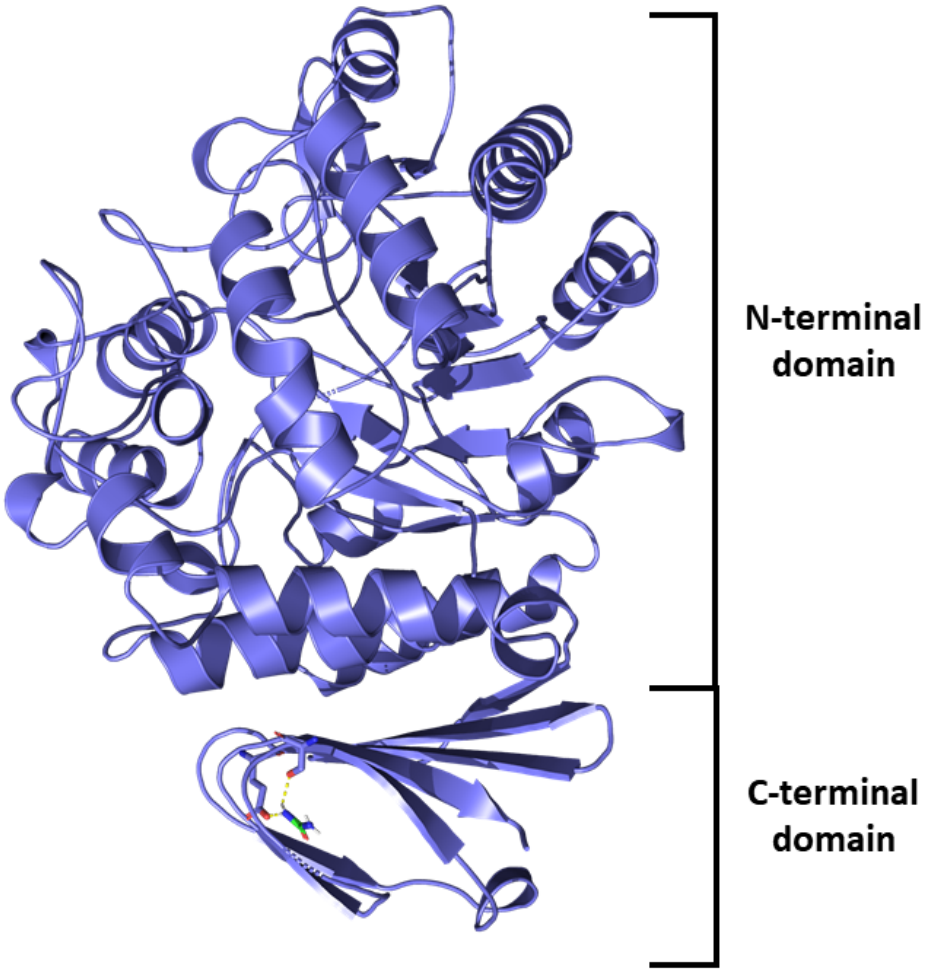

Figure 3.2. Structure of MalL crystallised with 1.5 M urea revealed a binding pocket in the C-terminal domain. Residues S536 and E554 with urea bound are shown as sticks (unpublished).

Based on the resulting crystal structures of urea bound to MalL, computational protein design (Section 1.3.2.2) was used by our collaborators at the Institute for Protein Innovation (Harvard) with the aim of designing an engineered version of MalL that would replicate the altered kinetics seen previously by the addition of urea. This computational technique resulted in the design of two MalL mutants where the enzyme was mutated to mimic the effect of urea binding. The two mutants developed comprised a single mutant where serine 536 was changed to an arginine (S536R) and a triple mutant where the aforementioned serine 536 was altered, in addition to the closely located glutamic acid 554 and valine 556 (S536D, E554Q, V556R).

3.2 MalL Wildtype (WT)

MalL WT was characterised as a baseline for comparison with all subsequent mutants and to ensure experimental conditions are appropriate.

3.2.1 Expression trials of MalL WT

MalL WT was expressed and purified using established protocols for MalL (Section 2.1) [Adapted from Schönert *et al.* (1998)]. MalL WT was expressed under the control of the *lac* operon in the pPROEX-Htb expression vector in *E. coli* DH5α. Pure soluble protein was obtained by immobilised metal affinity chromatography (IMAC) and size exclusion chromatography (SEC). Purity was confirmed by polyacrylamide gel electrophoresis (PAGE) (Section 2.1.4) (Figure 3.3).

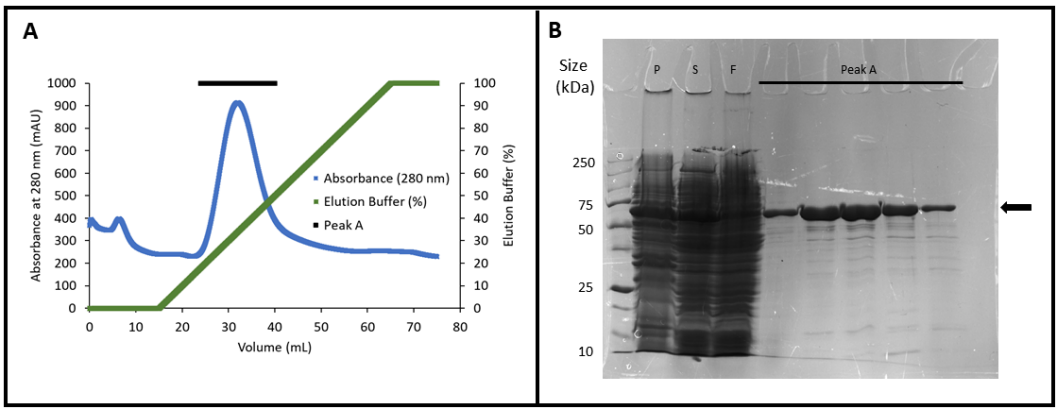

Figure 3.3. Ni IMAC purification of MalL WT. A) Elution chromatogram of MalL WT from IMAC column. B) PAGE of IMAC purification of MalL WT. P = Insoluble pellet, S = Supernatant containing protein loaded onto column, F = Flow-through of column minus bound target protein. Arrow indicates target protein.

MalL WT eluted from the IMAC column as a single peak at approximately 25% elution buffer. SDS-PAGE of fractions from under the peak indicates that it contained a mixed population of protein with MalL being the major constituent at approximately 90%. An intense band between 50 and 75 kDa is consistent with the expected MalL size of 69.5 kDa. MalL was further purified by SEC. MalL was eluted between 65 and 80 mL with improved purity. This corresponds to an estimated

molecular weight of 69.5 kDa which indicates MalL was purified as a soluble monomer. Purity was confirmed by SDS-PAGE (Figure 3.4).

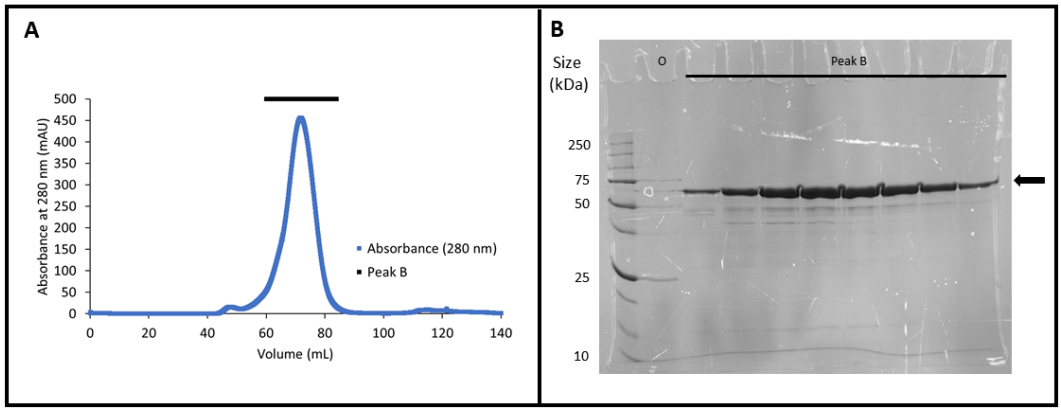

Figure 3.4. SEC purification of MalL WT. A) Elution chromatogram of MalL WT from SEC column. B) PAGE of SEC purification of MalL WT. O = Overflow from ladder well. Arrow indicates target protein.

3.2.2 Characterisation of MalL WT

Kinetic assays were performed based on established protocols for MalL (Section 2.2) with the substrate *p*-nitrophenyl-α-D-glucopyranoside (PNG).

3.2.2.1 Michaelis-Menten kinetics

MalL WT was characterised using the Michaelis-Menten kinetic model, to find the k_{cat}, K_M, and K_i (Figure 3.5). MalL was assayed at 25°C and this temperature was used for all subsequent Michaelis-Menten assays.

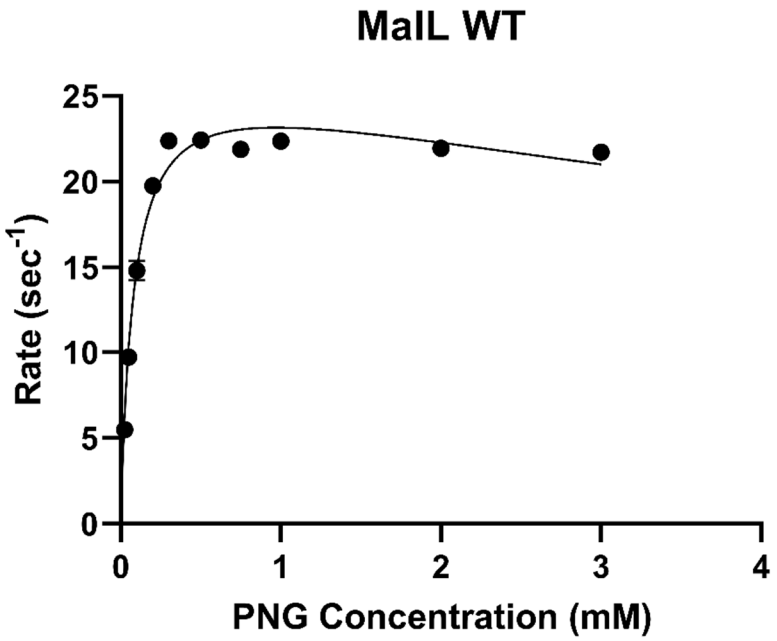

Figure 3.5. Michaelis-Menten kinetics of MalL WT at 25°C with the substrate p-nitrophenyl-α-D-glucopyranoside (PNG). Data fit with substrate-inhibition model. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

Michaelis-Menten parameters are given in Table 3.1. Data was fit with a substrate inhibition model with a K_i of 11.5 ± 1.8 mM. MalL was found to have a K_M of 0.0817 ± 0.005 mM. Subsequent assays were performed at approximately 10 X K_M at 1 mM substrate.

Table 3.1. Michaelis-Menten constants for MalL WT

	Enzyme
Parameter	**MalL WT**
k_{cat} (sec^{-1})	27.1 ± 0.6
K_M (mM)	0.0817 ± 0.005
K_i (mM)	11.5 ± 1.8

3.2.2.2 Temperature characterisation of MalL WT

The effect of temperature on the rate of MalL WT was assayed as described in section 2.2.2 and the data was fitted with a temperature dependant $\Delta C_P^{\ddagger}$ MMRT equation (Figure 3.6).

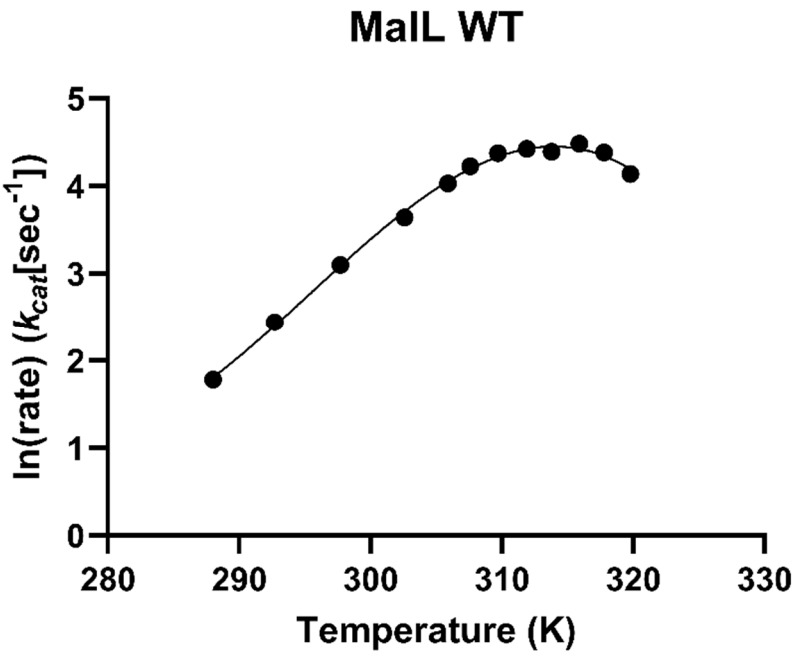

Figure 3.6. Results of temperature characterisation of MalL WT fit to the MMRT equation with a temperature dependant $\Delta C_P^{\ddagger}$. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

Results from the data from MalL WT are shown in Table 3.2. Results are consistent with an MMRT fit (R^2 = 0.9976). The optimum temperature (T_{opt}) and the heat capacity value ($\Delta C_P^{\ddagger}$) are of primary interest in the context of this study, however other variables such as T_0 (the reference temperature) are integral in implementing this equation. The change in enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) and the A value (the relationship of $\Delta C_P^{\ddagger}$ with temperature) are also variables of interest. MalL WT has a temperature optimum of 313.4 K (40.4°C). The heat capacity difference is -8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K).

Table 3.2. Parameters from the MMRT model fit to MalL WT data

	Enzyme
Parameter	**MalL WT**
T_0 **(K)**	309.4
$\Delta H^{\ddagger}$ **(**T_0**) (kJ mol^{-1})**	42.9 ± 1.2
$\Delta S^{\ddagger}$ **(**T_0**) (J mol^{-1} K^{-1})**	-70.7 ± 3.7
$\Delta C_P^{\ddagger}$ **(**T_0**) (kJ mol^{-1} K^{-1})**	-8.5 ± 0.3
A (kJ mol^{-1} K^{-2})	-219 ± 15
T_{opt} **(K)**	313.4

3.3 MalL S536R

Based on computational protein design the MalL mutant MalL S536R was proposed (Figure 3.7). MalL S536R is designed to mimic the effect of urea binding in the C-terminal domain by placing the side chain group of the introduced arginine into the binding pocket of urea.

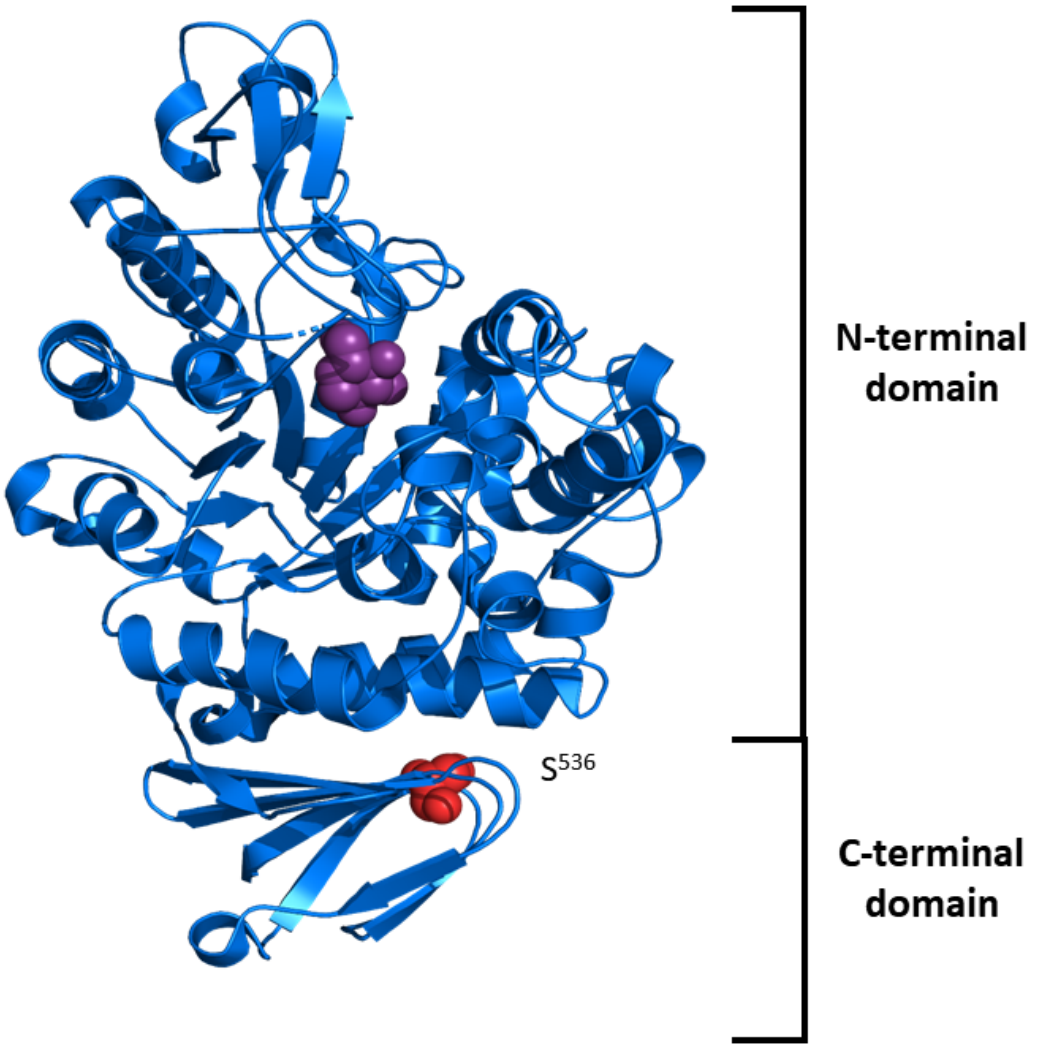

Figure 3.7. Position of S536R mutation in the MalL WT enzyme [PDB 5WCZ (van der Kamp et al., 2018)]. Mutated residue (S^{536}) is shown as red spheres. Substrate analogue (1-Deoxynojirimycin) is shown as purple spheres.

3.3.1 Expression trials of MalL S536R

Mutagenesis and cloning of the single mutant S536R was performed by E. Prentice using established "two-halves PCR" protocols. MalL S536R was expressed and purified according to established protocols for MalL variants (Section 2.1). Soluble protein was obtained by purification by immobilised metal ion affinity chromatography (IMAC) and size exclusion chromatography (SEC) with purity confirmed by polyacrylamide gel electrophoresis (PAGE).

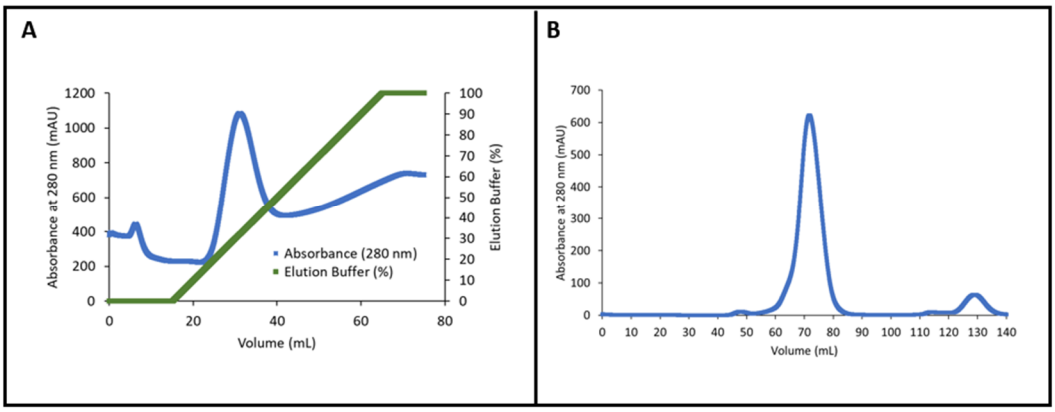

Figure 3.8. Purification of MalL S536R. A) Ni IMAC elution chromatogram. B) SEC elution chromatogram.

MalL S536R eluted from the IMAC column as a single peak at approximately 25% elution buffer. The peak contained MalL protein at approximately 90% purity as assessed by PAGE (gel not shown). MalL S536R was further purified by SEC. MalL S536R eluted from the column between 60 and 80 mL which corresponded to a molecular weight of 69.5 kDa in agreement with the purification of a soluble monomer of MalL. Purity was confirmed by PAGE (gel not shown).

3.3.2 Characterisation of MalL S536R

MalL S536R was assayed based on established protocols for MalL variants (Section 2.2) using the substrate *p*-nitrophenyl-α-D-glucopyranoside (PNG).

3.3.2.1 Michaelis-Menten kinetics

The Michaelis-Menten kinetic model was used to find the k_{cat}, K_M and K_i for MalL S536R (Figure 3.9).

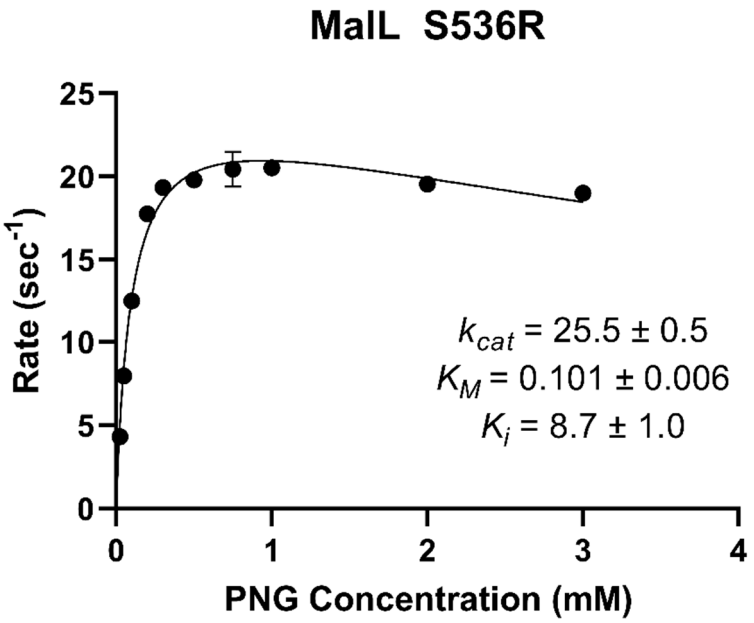

Figure 3.9. Michaelis-Menten kinetics of MalL S536R at 25°C with the substrate PNG (p-nitrophenyl-α-D-glucopyranoside). Data fit with substrate inhibition model. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

The results show that MalL S536R has a k_{cat} of 25.5 ± 0.5 sec^{-1} which is slightly less than that of the WT at 27.1 ± 0.6 sec^{-1} (Table 3.2). MalL S536R also shows more significant substrate inhibition compared to the WT as evidenced by its lower K_i value. Likewise, the K_M of S536R is slightly higher compared to the wildtype enzyme (Table 3.3).

Table 3.3. Michaelis-Menten constants for MalL S536R compared with MalL WT

	Enzyme	
Parameter	**MalL WT**	**MalL S536R**
k_{cat} **(sec^{-1})**	27.1 ± 0.6	25.5 ± 0.5
K_M **(mM)**	0.0817 ± 0.005	0.101 ± 0.006
K_i **(mM)**	11.5 ± 1.8	8.7 ± 1.0

3.3.2.2 Temperature characterisation of MalL S536R

The response of MalL S536R to temperature was measured as described in section 2.2.2. The data was fitted with a temperature dependant $\Delta C_P^{\ddagger}$ MMRT equation (Figure 3.10).

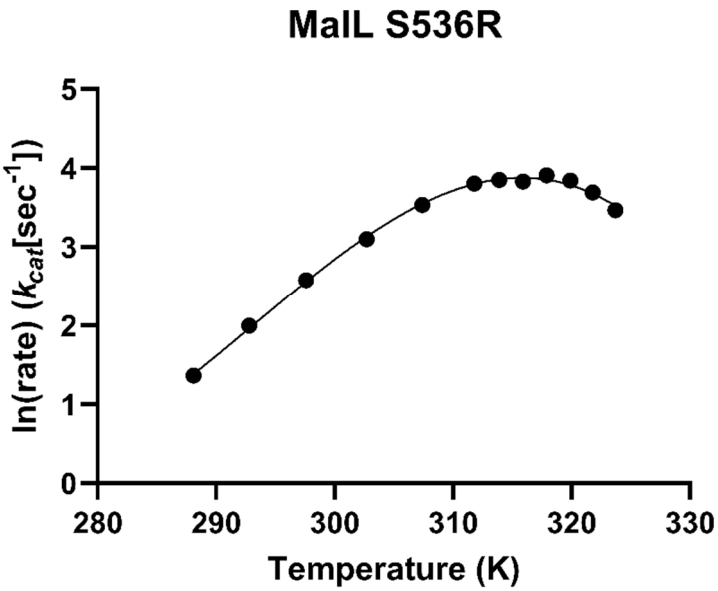

Figure 3.10. Temperature characterisation of MalL S536R fit with the temperature dependant $\Delta C_P^{\ddagger}$MMRT equation. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

MalL S536R shows a curved temperature dependence consistent with an MMRT model (R^2 = 0.9979) and has an increased temperature optimum of 316.8 K

(43.7 °C). Compared with WT, MalL S536R shows a significant drop in the rate across a range of temperatures when compared to the WT enzyme (Figure 3.11).

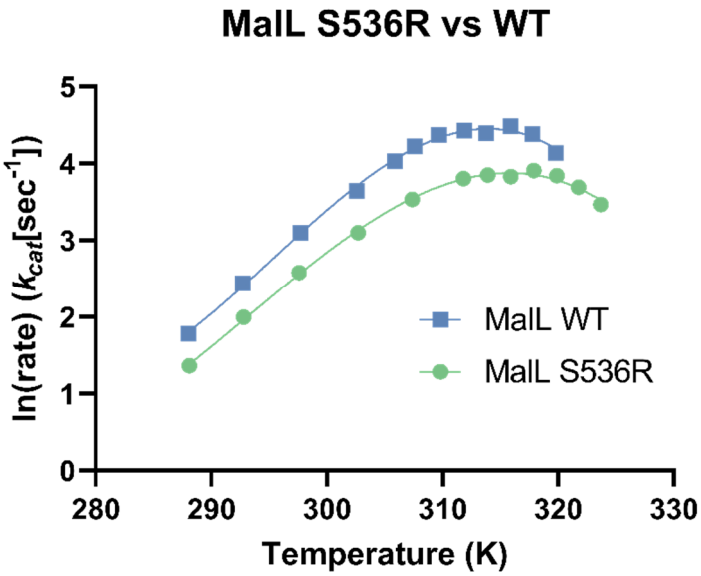

Figure 3.11. Temperature characterisation of MalL S536R versus MalL WT fit with the temperature dependant $\Delta C_P^{\ddagger}$MMRT equation.

MMRT fit parameters from the temperature characterisation data are shown in Table 3.4. Along with a drop in rate across a range of temperatures, MalL S536R also shows a significant increase in $\Delta C_P^{\ddagger}$ (-7.4 ± 0.2 kJ mol^{-1} K^{-1}) (T_0 = 312.8 K) compared to WT (-8.5 ± 0.3 kJ mol^{-1} K^{-1}) (T_0 = 309.4).

Table 3.4. MMRT fit parameters of MalL S536R and MalL WT

	Enzyme	
Parameter	**MalL WT**	**MalL S536R**
T_0 **(K)**	309.4	312.8
$\Delta H^{\ddagger}$ **(**T_0**) (kJ mol^{-1})**	42.9 ± 1.2	21.4 ± 0.8
$\Delta S^{\ddagger}$ **(**T_0**) (J mol^{-1} K^{-1})**	-70.7 ± 3.7	-145.1 ± 2.5
$\Delta C_P^{\ddagger}$ **(**T_0**) (kJ mol^{-1} K^{-1})**	-8.5 ± 0.3	-7.4 ± 0.2
A (kJ mol^{-1} K^{-2})	-219 ± 15	-139 ± 9
T_{opt} **(K)**	313.4	316.8

3.3.2.3 Characterisation of MalL WT with guanidinium hydrochloride

Urea has been shown to dynamically effect the enzymatic rate of MalL (Section 3.1). Similar dynamic effects were observed in the MalL S536R mutant which mimicked the urea binding. We sought to replicate the effects of urea with another denaturant, guanidinium hydrochloride. Urea, arginine and guanidinium share similar functional groups (Figure 3.12).

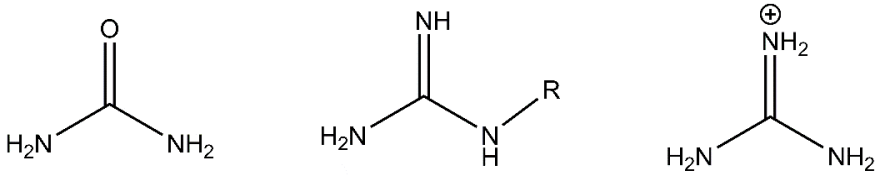

Figure 3.12. Similarity in functional groups of urea, arginine and guanidinium (Left to right, respectively).

MalL WT was assayed for temperature dependence with the addition of guanidinium hydrochloride (GuHCl) based on a protocol used by Parker *et al*., (1995). In their study, the N-terminal domain of phosphoglycerate kinase showed unfolding between 0.8 and 1.2 M GuHCl (Parker *et al.*, 1995). MalL was therefore assayed with 1 M GuHCl as an initial assay concentration, however at this concentration MalL activity was found to be completely abolished. The assay was then repeated with 0.1 M GuHCl initially over a small temperature range (Figure

3.13), and subsequently a full temperature characterisation was collected (Figure 3.14).

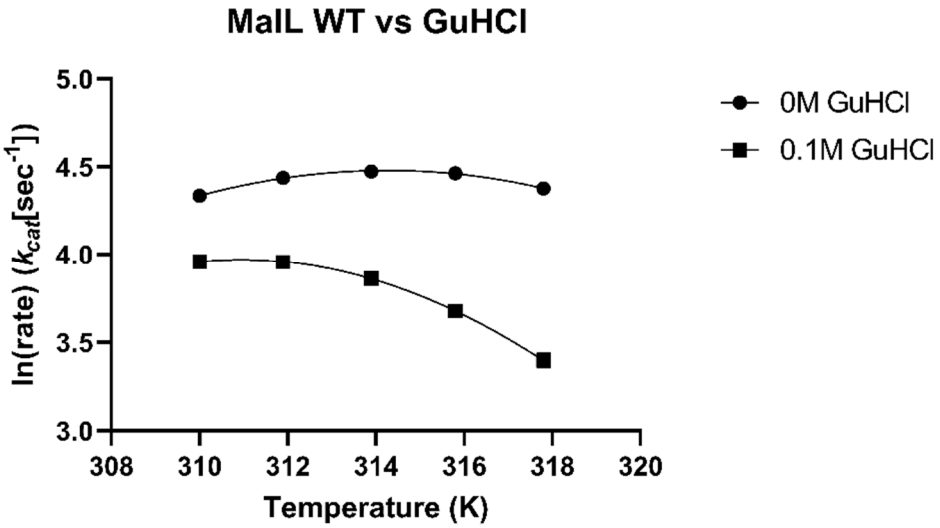

Figure 3.13. Initial narrow temperature range characterisation of MalL WT with 0.1M GuHCl fit with the MMRT equation using a temperature dependent $\Delta C_P^{\ddagger}$. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

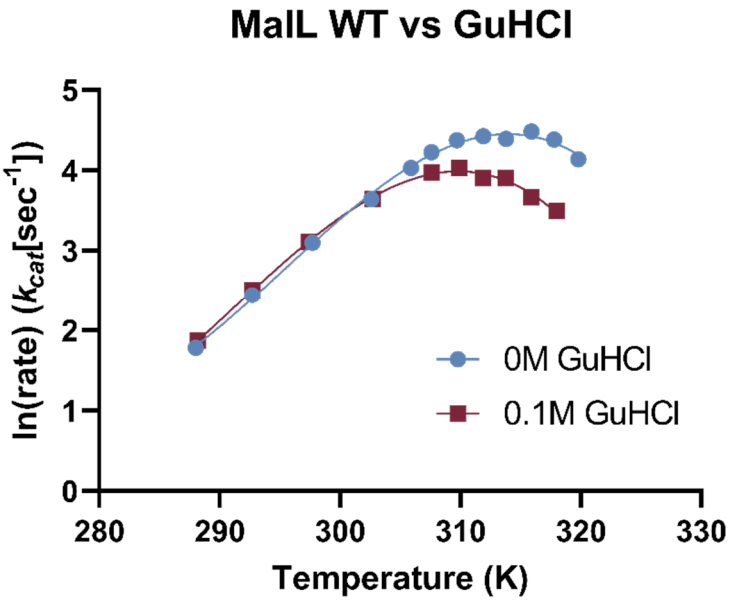

Figure 3.14. Temperature characterisation of MalL WT with 0.1M GuHCl fit with the MMRT equation using a temperature dependent $\Delta C_P^{\ddagger}$. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

The results from these assays show that GuHCl appears to have little effect on the rate of MalL until approximately 303 K (30°C) where it deviates from the WT curvature, resulting in a lowered temperature optimum of 309.7 K (36.6°C). The $\Delta C_P^{\ddagger}$ is unchanged with the addition of GuHCl at -8.5 ± 0.2 kJ mol^{-1} K^{-1} (T_0 = 305.7 K), compared to -8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K) without GuHCl (Table 3.5).

Table 3.5. MMRT fit parameters of MalL WT with GuHCl

	Enzyme	
Parameter	**0 M GuHCl**	**0.1 M GuHCl**
T_0 (K)	309.4	305.7
$\Delta H^{\ddagger}$ (T_0) (kJ mol^{-1})	42.9 ± 1.2	38.3 ± 1.0
$\Delta S^{\ddagger}$ (T_0) (J mol^{-1} K^{-1})	-70.7 ± 3.7	-87.7 ± 3.4
$\Delta C_P^{\ddagger}$ (T_0) (kJ mol^{-1} K^{-1})	-8.5 ± 0.3	-8.5 ± 0.2
A (kJ mol^{-1} K^{-2})	-219 ± 15	-211 ± 13
T_{opt} (K)	313.4	309.7

3.4 MalL Triple Mutant

3.4.1 Cloning of MalL triple mutant

The MalL triple mutant (TM) was designed to mimic the effect of urea binding in the C-terminal domain. The MalL TM gene construct was ordered from GeneArt (ThermoFisher scientific) codon optimised for *E. coli*, in the pMA-T cloning vector with ampicillin (AMP) resistance. The process of manipulating this construct for recombinant protein expression was briefly as follows. MalL TM in pMA-T was transformed into *E. coli* and successful transformants selected by growth on LB AMP agar plates. Selected colonies were grown in 5 mL AMP LB cultures to isolate plasmid DNA (Section 2.3.7.2). MalL TM was digested from the plasmid (Section 2.3.3) and the MalL TM gene (1695 bp) isolated by agarose gel electrophoresis (Figure 3.15).

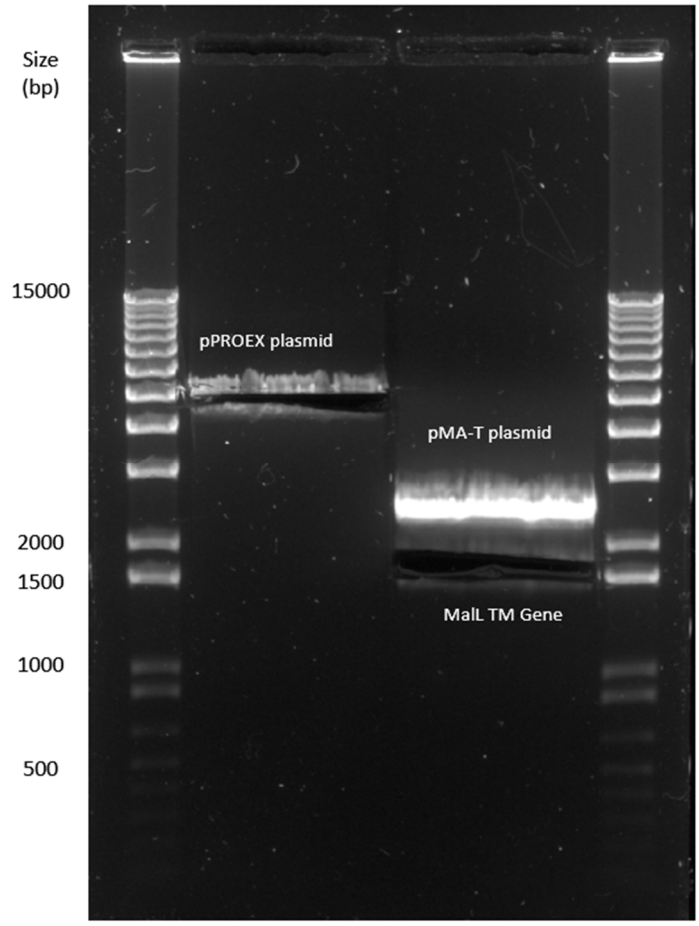

Figure 3.15. Restriction enzyme digest of MalL TM from pMA-T cloning vector and digest of pPROEX-Htb expression vector. Run on a 1% TAE agarose gel. Bands of interest were excised from the gel as shown here.

The MalL TM gene was then ligated into the pPROEX-Htb expression vector (Section 2.3.4) and transformed (Section 2.3.5) into *E. coli* DH5α. Transformed *E. coli* were grown on LB AMP agar plates and selected colonies screened for the 1852 bp gene insert by colony PCR (Section 2.3.6) (Figure 3.16).

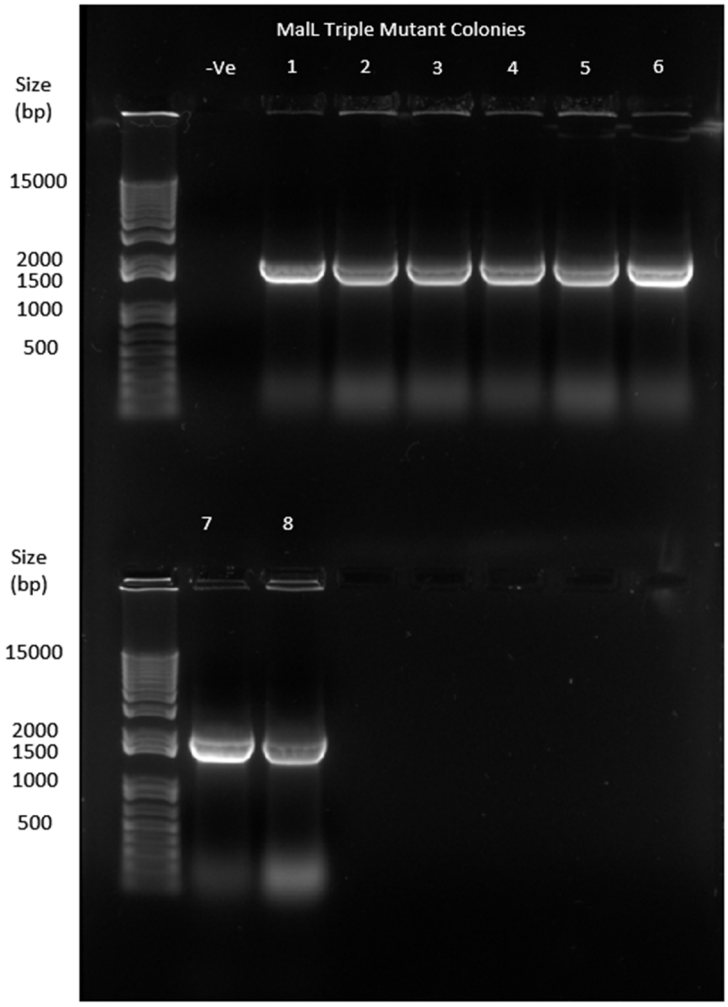

Figure 3.16. Agarose gel products of a colony PCR of the MalL TM gene. Eight colonies were selected from the transformation of MalL TM into E. coli. Run on a 1% TAE agarose gel.

All colonies had the expected insert size. Two colonies were selected and sent for sequencing (Section 2.3.8). Sequencing results showed both colonies contained intact inserts of MalL TM. MalL TM with the mutations (S536D, E554Q, V556R) is shown in Figure 3.17.

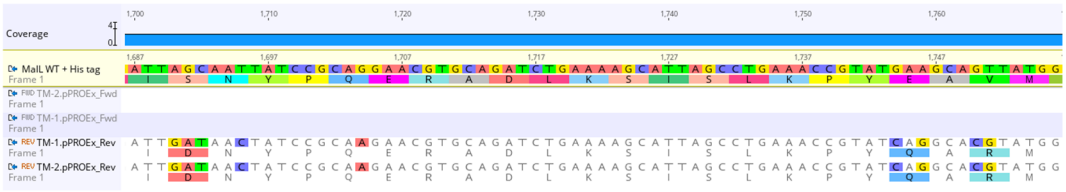

Figure 3.17. Sequencing results of MalL TM shows successful insert of gene construct.

3.4.2 Expression trials of MalL triple mutant

MalL TM was expressed and purified according to established protocols for MalL variants (Section 2.1). MalL TM was expressed in *E. coli* in the pRPOEX-Htb expression vector and induced by IPTG induction. Soluble protein was obtained by purification by immobilised metal ion affinity chromatography (IMAC) (Figure 3.18) and its purity confirmed by polyacrylamide gel electrophoresis (PAGE). Further protein purity was achieved by size exclusion chromatography (SEC) (Figure 3.19).

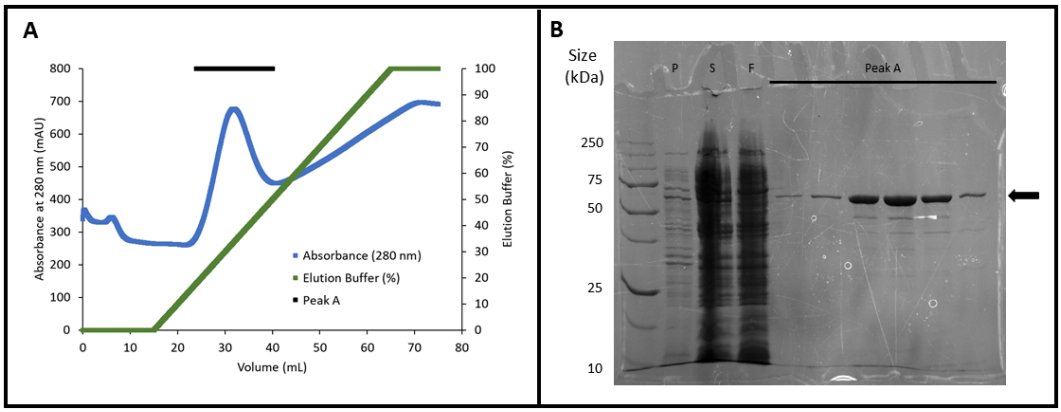

Figure 3.18. Ni IMAC Purification of MalL TM. A) Elution chromatogram of MalL TM. B) PAGE gel of MalL TM IMAC purification. P = Insoluble pellet, S = Supernatant containing protein loaded onto column, F = Flow-through from column minus bound target protein. Arrow indicates target protein.

MalL TM was eluted as a single peak at approximately 20% elution buffer from the Ni IMAC column. The peak had a majority population of MalL identified by denaturing PAGE. PAGE showed an intense band between 50 and 75 kDa corresponding to the expected MalL size of 69.5 kDa (as shown by the arrow). MalL TM was additionally purified by SEC with improved purity confirmed by PAGE. MalL TM eluted between 60 and 80 mL which corresponds to a molecular weight of 69.5 kDa indicating MalL was purified as a soluble monomer (Figure 3.19).

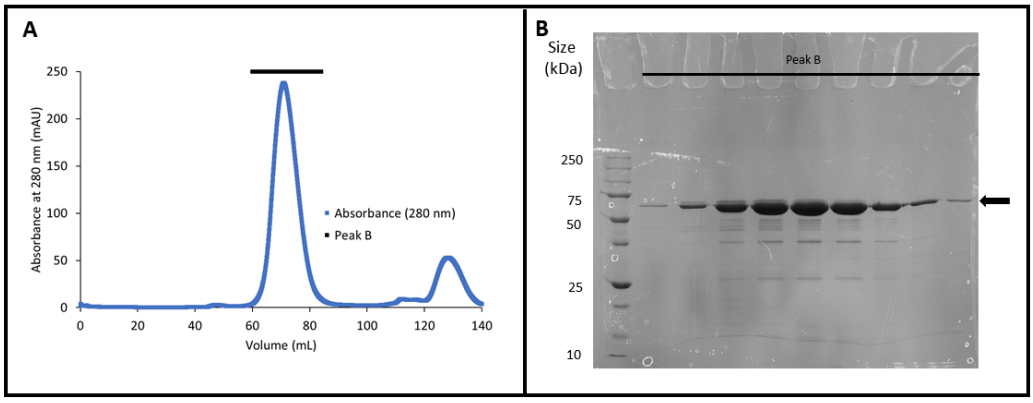

Figure 3.19. SEC purification of MalL TM. A) Chromatogram of MalL TM SEC elution. B) PAGE gel of MalL TM SEC purification. Arrow indicates target protein.

3.4.3 Characterisation of MalL triple mutant

Kinetic assays were performed using established protocols for MalL variants (Section 2.2) using the substrate *p*-nitrophenyl-α-D-glucopyranoside (PNG).

3.4.3.1 Michaelis-Menten kinetics

The Michaelis-Menten kinetic model was used to find the k_{cat}, K_M, and K_i for MalL TM. The results are shown in Figure 1.21.

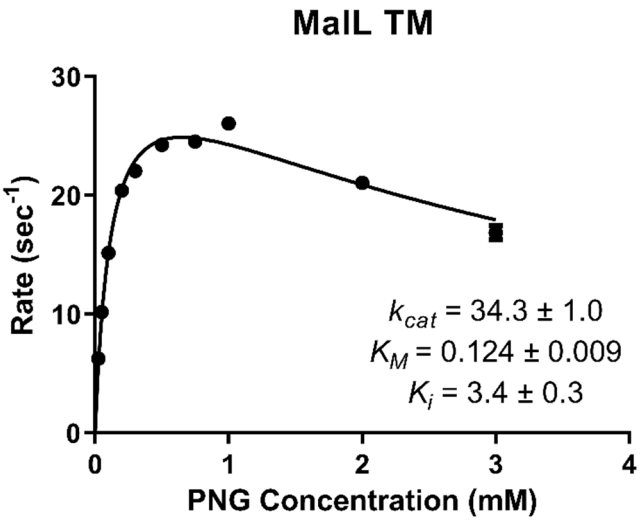

Figure 3.20. Michaelis-Menten kinetics of MalL TM. Data is fitted with a substrate inhibition model. Substrate p-nitrophenyl-α-D-glucopyranoside (PNG). Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

MalL TM has a k_{cat} of 34.3 ± 1.0 sec^{-1} and a K_M of 0.124 ± 0.009 mM which are both higher than MalL WT at 27.1 ± 0.6 sec^{-1} and 0.0817 ± 0.005 mM, respectively (Table 3.6). Substrate inhibition is much more pronounced in MalL TM than WT with a K_i of only 3.4 ± 0.3 mM compared with 11.5 ± 1.8 mM (Figure 3.21).

Table 3.6. Michaelis-Menten constants for MalL TM compared with MalL WT.

	Enzyme	
Parameter	**MalL WT**	**MalL TM**
k_{cat} (sec^{-1})	27.1 ± 0.6	34.3 ± 1.0
K_M (mM)	0.0817 ± 0.005	0.124 ± 0.009
K_i (mM)	11.5 ± 1.8	3.4 ± 0.3

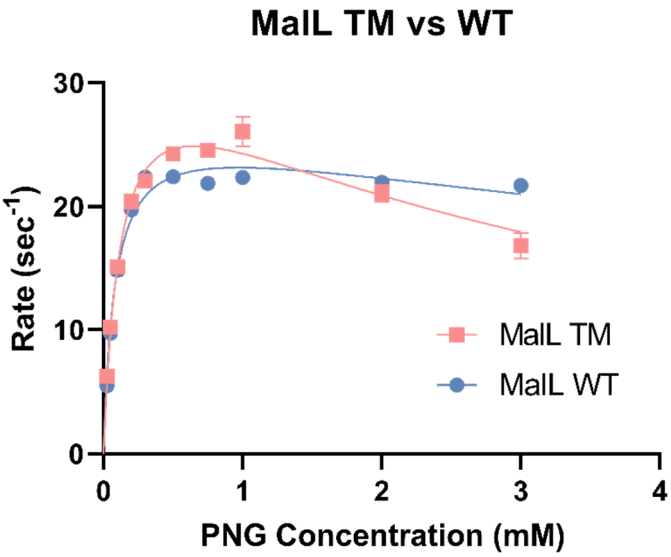

Figure 3.21. Michaelis-Menten kinetics of MalL TM versus MalL WT.

3.4.3.2 Temperature characterisation of MalL triple mutant

The activity of MalL TM was assayed with temperature using normal protocols for MalL (Section 2.2.2). Data was fitted to a temperature dependent $\Delta C_P^{\ddagger}$ MMRT equation (Figure 3.22).

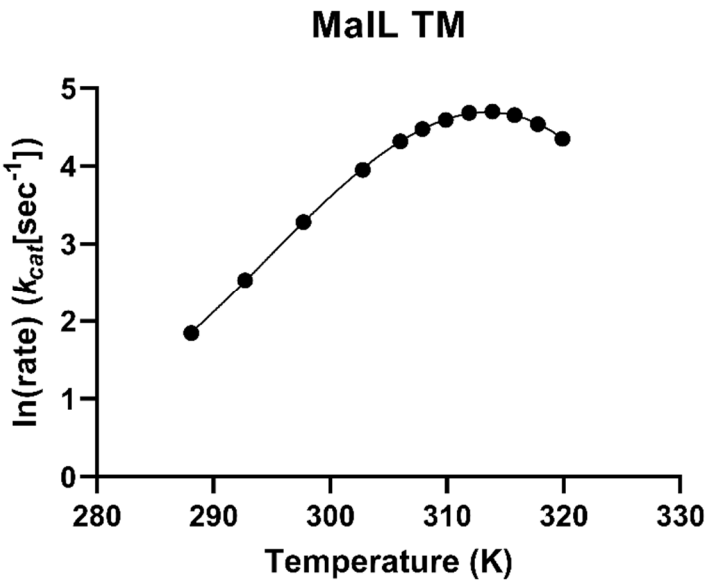

Figure 3.22. Temperature characterisation of MalL TM fit with the temperature dependant $\Delta C_P^{\ddagger}$ MMRT model. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

MalL TM shows a curved temperature dependence consistent with MMRT (R_2 = 0.9996). MalL TM has a temperature optimum of 313.1 K (40°C) which is almost identical to MalL WT. MalL TM shows a reasonably similar temperature profile to that of the WT enzyme (Figure 3.23), however it has a slightly higher rate overall with a $\Delta C_P^{\ddagger}$ of -9.3 ± 0.1 (T_0 = 309.1 K), which is slightly lower than that of MalL WT (Table 3.7).

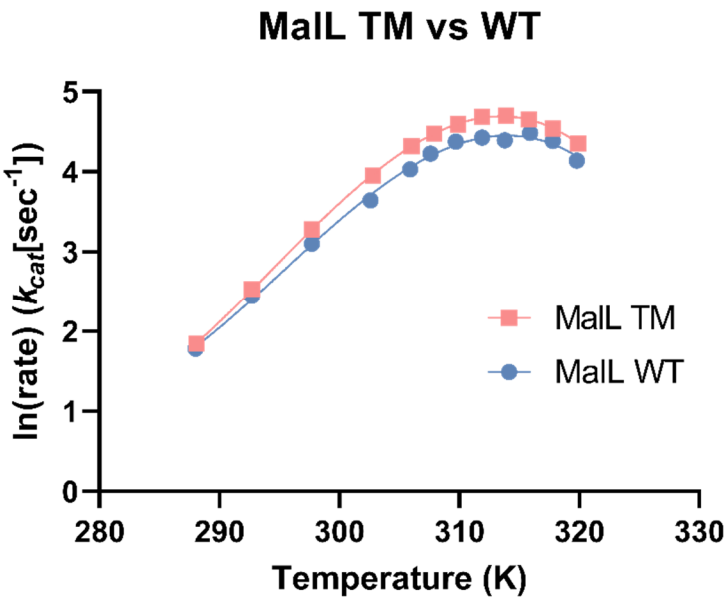

Figure 3.23. Temperature characterisation of MalL TM versus WT fit with the temperature dependant $\Delta C_P^{\ddagger}$ MMRT model.

Table 3.7. MMRT fit parameters of MalL TM and MalL WT

	Enzyme	
Parameter	**MalL WT**	**MalL TM**
T_0 (K)	309.4	309.1
$\Delta H^{\ddagger}$ (T_0) (kJ mol^{-1})	42.9 ± 1.2	44.9 ± 0.4
$\Delta S^{\ddagger}$ (T_0) (J mol^{-1} K^{-1})	-70.7 ± 3.7	-62.1 ± 1.4
$\Delta C_P^{\ddagger}$ (T_0) (kJ mol^{-1} K^{-1})	-8.5 ± 0.3	-9.3 ± 0.1
A (kJ mol^{-1} K^{-2})	-218 ± 15	-229 ± 6
T_{opt} (K)	313.4	313.1

3.5 Discussion

MalL and its variants have provided the platform to explore the effects of allostery. We have successfully shown that mutations far away from the active site can have significant effects on enzymatic rates, via changes in enzyme dynamics.

MalL S536R achieved a significant decrease in the rate over the WT enzyme with a k_{cat} of 25.5 ± 0.5 sec^{-1} compared to WT at 27.1 ± 0.6 sec^{-1} at 25°C. The S536R

variant resulted in an increased temperature optimum and $\Delta C_P^{\ddagger}$ compared to the WT enzyme, with a T_{opt} of 316.8 K (43.7 °C) and a $\Delta C_P^{\ddagger}$ of -7.4 ± 0.2 kJ mol^{-1} (T_0 = 312.8 K) compared to a T_{opt} of 313.4 K (38.3°C) and a $\Delta C_P^{\ddagger}$ -8.5 ± 0.3 kJ mol^{-1} (T_0 = 309.4 K) in MalL WT. This is consistent with a rigidification of the C-terminal domain, which is shown by predicted stability changes (calculated by Normal Mode Analysis) (Section 2.5), in Figure 3.24 (Rodrigues *et al.*, 2018).

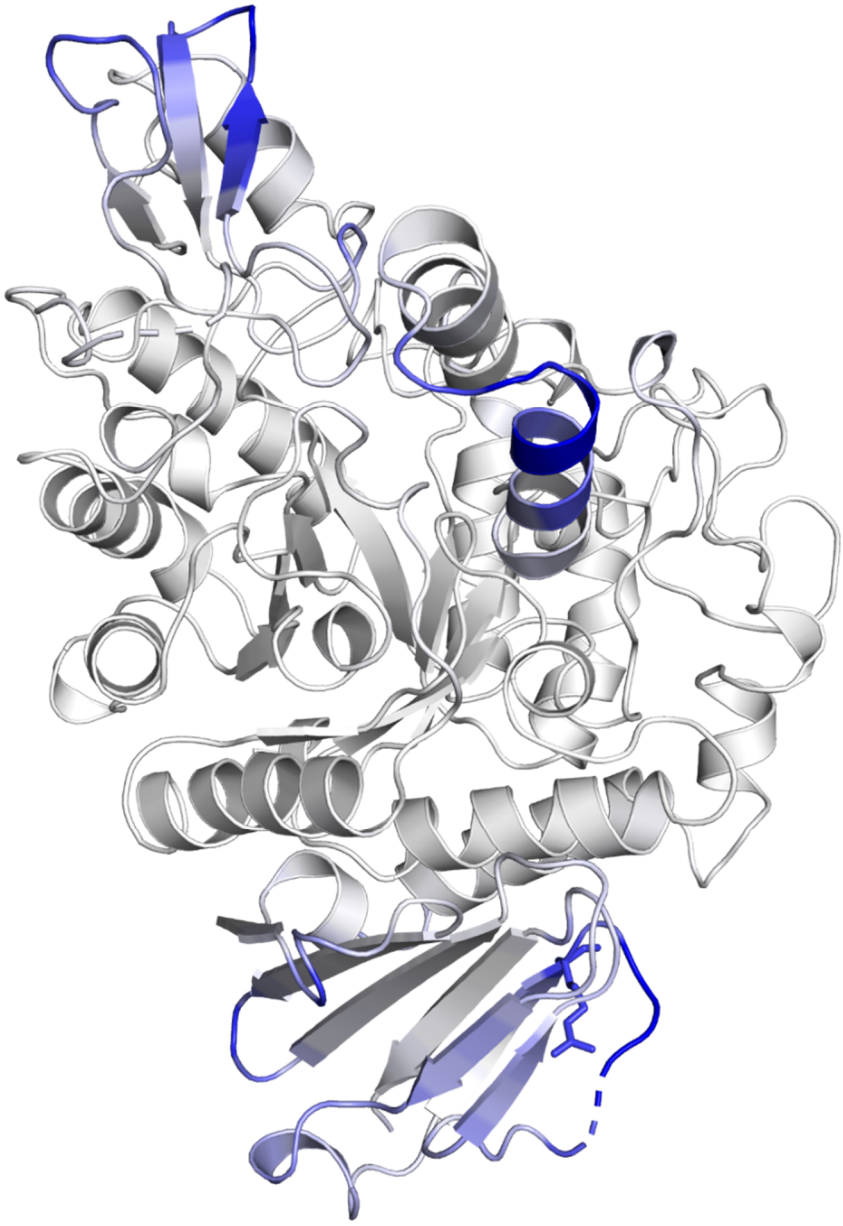

Figure 3.24. Predicted dynamic changes in MalL S536R based on normal mode analysis. Blue regions are regions of increased rigidity compared to the WT enzyme.

MalL S536R was able to exert its effect on the rate from approximately 30 Å away from the active site. This shows that allosteric effects can be transmitted across domain boundaries to effect catalytic rate via heat capacity changes. The Michaelis-Menten kinetics indicate a much subtler effect on the enzymatic rate than that observed in the temperature characterisation. We can make use of the A value from the MMRT equation, which is the relationship of heat capacity with temperature. A large negative number indicates that $\Delta C_P^{\ddagger}$ is negatively correlated with an increase in temperature. Interestingly for S536R, the $\Delta C_P^{\ddagger}$ is not as strongly correlated with temperature as MalL WT, as the A value is only -139 ± 9 kJ mol^{-1} K^{-2} for S536R compared to -219 ± 15 kJ mol^{-1} K^{-2} in WT. This means that as temperature increases, the value for $\Delta C_P^{\ddagger}$ for MalL S536R will not decline as much as WT, in other words it will remain more rigid.

Previous research (Hobbs *et al*., 2013) showed a relationship between two kinetic terms, with an increase in T_{opt} correlating to a less negative $\Delta C_P^{\ddagger}$ value. With MalL S536R, we have seen results consistent with this; T_{opt} increased and the $\Delta C_P^{\ddagger}$ value increased. In other words, the enzyme became less dynamic but was able to increase the temperature at which it was optimal.

Addition of GuHCl was used to mimic the effect of the arginine mutation in the C-terminal domain. Although a strong deviation at elevated temperatures was observed in the temperature characterisation with the addition of 0.1 M GuHCl, overall parameter fit of the MMRT equation is very similar. The $\Delta C_P^{\ddagger}$ with the addition of 0.1 M GuHCl [-8.5 ± 0.2 kJ mol^{-1} K^{-1} (T_0 = 305.7 K)] is within margin of error of MalL WT [-8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K)]. The addition of 0.1 M GuHCl produces a similar profile to that of urea with slight activation at low temperatures and inhibition at higher temperatures with a decreased T_{opt}.

MalL TM has both an increased k_{cat} and increased substrate inhibition (lower K_i). The $\Delta C_P^{\ddagger}$ in MalL TM has increased compared to the WT enzyme. The $\Delta C_P^{\ddagger}$ was -8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K) in the WT enzyme compared to -9.3 ± 0.1 kJ mol^{-1} K^{-1} (T_0 = 309.1 K) for TM. This seems to indicate that in contrast to the S536R mutant the triple mutation has had an overall destabilising

effect on the MalL enzyme. The increased rate for this enzyme may have been achieved via enzyme-substrate destabilisation.

Crystal structures of MalL S536R and TM are presented in Section 4. In the future these structures would allow for the analysis of heat capacity changes in these mutants by molecular dynamics simulations. Molecular dynamics simulations can allow for visualisation of molecular motions over time. This allows for observations of how flexibility and heat capacity is changing across the enzyme. This would allow for a better understanding of how these mutations exert their effects on the enzymatic rate. In addition, there is some disagreement between the Michaelis-Menten and temperature characterisation that the cause of which would need to be resolved.

4 Crystal structures of MalL mutants S536R and the triple mutant (S536D, E554Q, V556R)

4.1 Introduction

The MalL S536R and MalL triple mutant (TM) were created to mimic the effects of urea binding in the C-terminal domain of MalL. These two mutants were characterised kinetically, and the results showed significant changes to their enzymatic rates (Sections 3.3.2 and 3.4.3). Following the kinetic characterisation, we sought to crystallise the mutant enzymes in order to characterise them structurally.

4.2 MalL S536R crystallisation and structure determination

4.2.1 Crystallisation

Initial crystallisation conditions were determined by screening from a library of 384 crystallisation conditions (Hampton Research) using sitting drop techniques (Section 2.4.1). Promising conditions were further screened by using hanging-drop techniques by varying ionic strength, pH, precipitant, drop size and protein concentration (Section 2.4.2). Final crystallisation conditions for MalL S536R were 0.1 M Tris pH 8.0, 0.2 M ammonium acetate and 18% w/v PEG 10,000. Protein at a concentration of 10 $mg.ml^{-1}$ was mixed at a 1:1 ratio with precipitant solution.

4.2.2 Data collection

Crystals were frozen in liquid nitrogen with 20% v/v glycerol added to the precipitant solution as a cryoprotectant. Crystals were taken to the Australian Synchrotron for data collection. Data were collected at a distance of 125 mm. Data were collected for a full 360° rotation in one degree increments for one second

per frame. Representative crystals and diffraction for MalL S536R are shown in Figure 4.1.

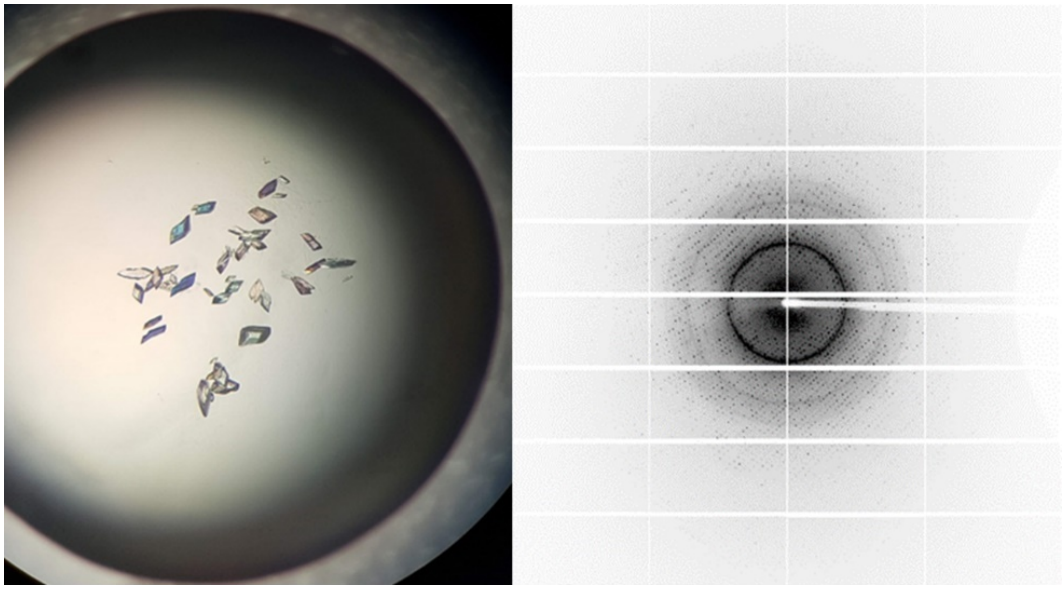

Figure 4.1. Representative crystals and X-ray diffraction for MalL S536R. X-ray diffraction recorded to 1.1Å.

4.2.3 Data processing

Data processing was performed in the CCP4 software suite (Winn *et al.*, 2011). Space group determination was performed in Pointless (Evans, 2011). Data was scaled and merged in AIMLESS (Evans & Murshudov, 2013). Matthews coefficient analysis was used to find the number of molecules in the asymmetric unit (Kantardjieff & Rupp, 2003; Matthews, 1968). Statistics from these processes is given in Table 4.1.

Table 4.1. Data collection statistics for MalL S536R.

Data Statistic	MalL S536R
Space group	P 1 2_1 1
Wavelength (Å)	0.9537
Cell Dimensions: **a/b/c (Å)** **α/β/γ (°)**	 48.69/100.95/61.72 90.0/113.1/90.0
Mosaicity	0.06
Monomers in the asymmetric unit	1
Resolution range (Å)	23.06-1.10 (1.12-1.10)
Number of observed reflections	1477816 (65945)
Number of unique reflections	208284 (9598)
R_{merge}	0.06 (0.51)
Mean I/σI	15.2 (3.5)
Completeness	94.1 (87.8)
Multiplicity	7.1 (6.9)
Wilson B-factor	9.79

* Values in brackets correspond to the outer resolution shell.

Molecular replacement using the previously solved MalL WT structure (PDB: 4M56) was performed using Molrep (Vagin & Teplyakov, 1997). Models were manually refined in COOT (Emsley *et al.*, 2010). Refinement of the model was performed using Refmac5 in CCP4 or Phenix (Afonine *et al.*, 2012; Murshudov *et al.*, 1997). A representative electron density map and model fit is shown in Figure 4.2.

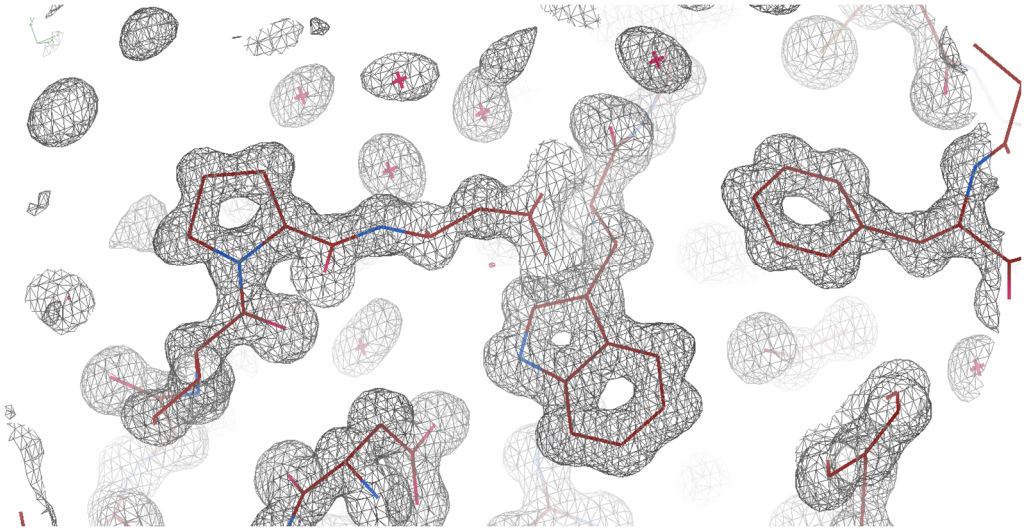

Figure 4.2. Representative electron density map and model fit of MalL S536R at 1.10 Å resolution. Contour level is set to 1.00 rmsd.

Thirty amino-acids were missing from the N-terminus including the hexa-His tag and were not modelled in the final structure (Full length protein sequence is included in Section 7.1). A small number of sidechains were unable to be modelled into the final density. Three additional molecules were identified in the crystal structure. Tris and glycerol from crystallisation conditions and cryoprotectant were identified near the active site. A calcium ion was found to be bound to Asp20, Asn22, Asp24, Phe26 and Asp28 which is consistent with calcium binding in other glycoside hydrolase enzymes. Refinement statistics are given in Table 4.2.

Table 4.2. Refinement statistics for MalL S563R.

Refinement statistic	MalL S536R
R-factor	0.127
R-free	0.145
Total number of atoms	5209
Total number of protein atoms	4713
Other molecules	3
Number of waters	481
RMSD:	
Bond length (Å)	0.0192
Bond angles (°)	2.0657
Average B-factors (Å^2)	
Protein monomer	16.16
Waters	24.76
Ramachandran analysis:	
Percentage in favoured regions	96.2
Percentage in allowed regions	3
Percentage in disallowed regions	0.8

4.3 MalL triple mutant (S536D, E554Q, V556R) crystallisation and structure determination

4.3.1 Crystallisation

Initial crystallisation conditions were determined by screening from a library of 288 conditions (Hampton Research) using sitting drop techniques. Promising conditions were further screened by using hanging-drop techniques where ionic strength, pH, precipitant and protein concentration was varied. Final crystallisation conditions for the triple mutant were 0.1 M Tris HCl pH 8.5, 0.22 M Sodium Acetate, 27% w/v PEG 4,000.

4.3.2 Data collection

Crystals were frozen in liquid nitrogen and taken to the Australian synchrotron for data collection. Crystals were frozen with 20% v/v glycerol added to the precipitant solution as a cryoprotectant. Data was collected for a full 360° rotation in one degree per second increments at a detector distance of 225 mm. Representative crystals and diffraction for MalL TM is shown in Figure 4.3.

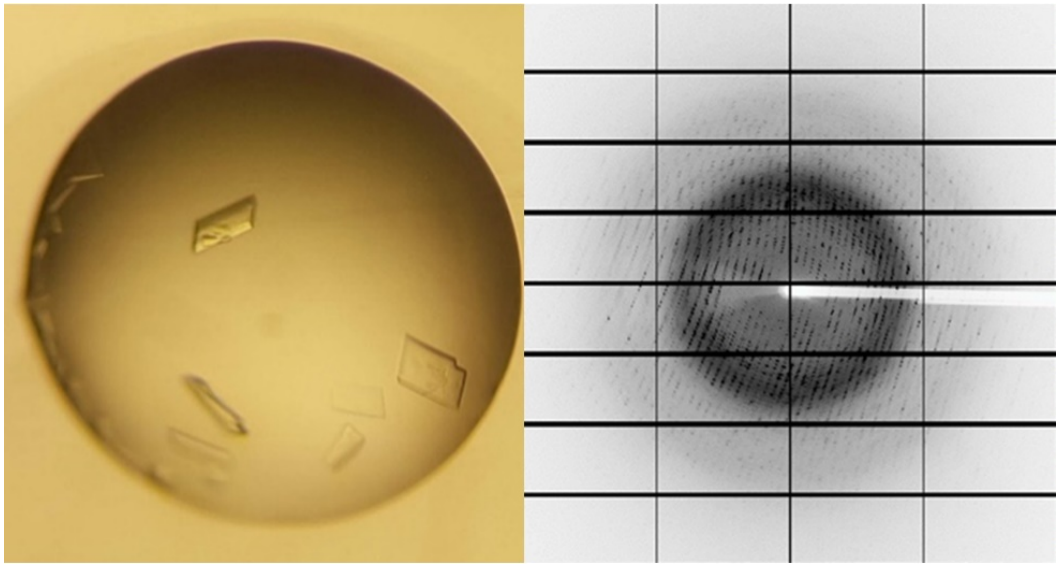

Figure 4.3. Representative crystals and X-ray diffraction for MalL TM. X-ray diffraction recorded to 1.6 Å.

4.3.3 Data processing

Data processing was performed using the CCP4 and PHENIX software suites (Adams *et al.*, 2010; Winn *et al.*, 2011). Space group determination was done using Pointless (Evans, 2011). Data was scaled and merged in AIMLESS (Evans & Murshudov, 2013). Data was cut off at this stage to 1.7 Å resolution to make I/σI approximately two in the outer resolution shell. Matthews coefficient analysis was performed to determine the number of monomers in the asymmetric unit (Kantardjieff & Rupp, 2003; Matthews, 1968). Data collection statistics are shown in Table 4.3.

Table 4.3. Data Statistics for MalL Triple Mutant.

Data Statistic	MalL TM
Space group	P 1 2_1 1
Wavelength (Å)	0.9537
Cell Dimensions	
a/b/c (Å)	49.40/89.39/59.00
α/β/γ (°)	90.0/113.1/90.0
Mosaicity	0.16
Monomers in the asymmetric unit	1
Resolution range (Å)	45.81-1.70 (1.73-1.70)
Number of observed reflections	360320 (19103)
Number of unique reflections	52199 (2731)
R_{merge}	0.08 (0.834)
Mean I/σI	12.5 (2.2)
Completeness	100 (100)
Multiplicity	6.9 (7)
Wilson B-factors	18.857

* Values in brackets correspond to the outer resolution shell.

Molecular replacement was performed using Phenix (Phaser-MR) using the previously solved MalL WT structure (PDB: 4M56) as a model (McCoy *et al.*, 2007). Manual model building was performed in COOT (Emsley *et al.*, 2010). Refinement was completed using Refmac5 and Phenix (Afonine *et al.*, 2012; Murshudov *et al.*, 1997). Representative electron density map and model fit is shown in Figure 4.4.

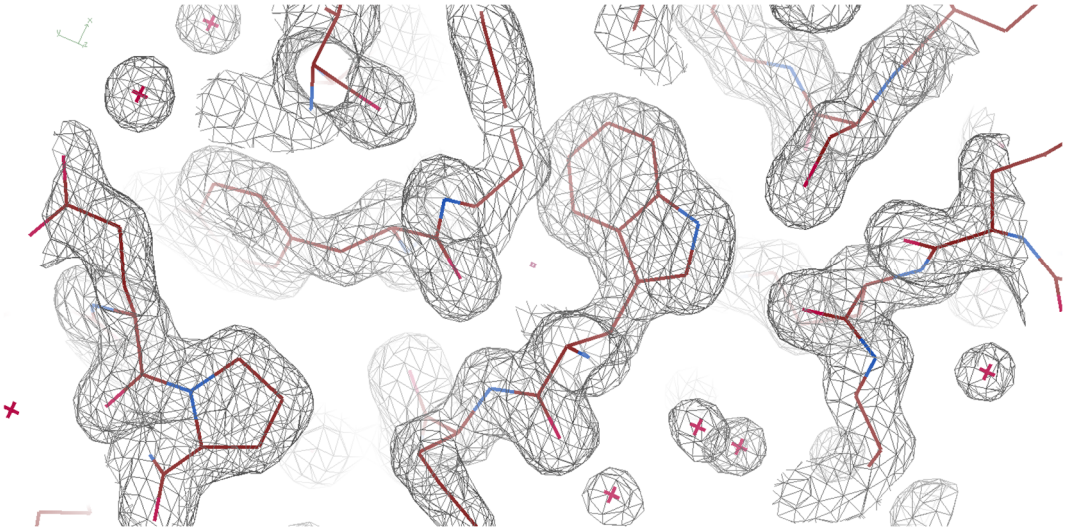

Figure 4.4. Electron density map and model quality fit of MalL TM at 1.7 Å resolution. Contour is set to 1.00 rmsd.

There were 29 amino-acids including the hexa-his tag at the C-terminus that were not modelled in the final structure (Full length protein sequence is included in Section 7.1). A small number of side chains throughout the structure were missing from the electron density. Side chains from these residues were not modelled in the final structure. Three additional molecules were identified in the crystal structure. Tris from the crystallisation conditions and glycerol from the cryoprotectant were identified near the active site. A calcium metal ion was identified near the C-terminus coordinated with the amino acids Asp20, Asn22, Asp24, Phe26 and Asp28. This calcium binding site is consistent with other MalL structures, and that of glycoside hydrolase group 13 (GH13) members in general. Refinement statistics for MalL TM are shown in Table 4.4.

Table 4.4. Refinement statistics for MalL TM.

Refinement statistic	MalL TM
R-factor	0.152
R-free	0.188
Total number of atoms	4910
Total number of protein atoms	4565
Other molecules	3
Number of waters	330
RMSD:	
Bond length (Å)	0.0109
Bond angles (°)	1.6268
Average B-factors (Å^2)	
Protein monomer	23.95
Waters	30.33
Ramachandran analysis:	
Percentage in favoured regions	94.6
Percentage in allowed regions	4
Percentage in disallowed regions	1.1

4.4 Discussion

4.4.1 Structural Comparison

The structures of MalL S536R and MalL TM show the same structural architecture as MalL WT, which is also seen in other GH13 enzymes. A comparison of all three structures viewed from one aspect highlights their similarity (Figure 4.5). Global Root Mean Square Deviation (RMSD) values (Calculated from C_α atoms) between MalL variants are between 0.3 and 0.7 Å, indicating a high level of structural similarity (Table 4.5). The unit cell parameters of MalL TM were nearly identical to that of MalL S536R, however had a 10% decrease in the unit cell size (100.95 to 89.39 Å). In contrast to the WT structure (PDB: 4M56) both the TM and S536R structures contained only one monomer in the asymmetric unit.

Mutation positions are approximately 30 Å away from the active site as illustrated by arrows/dots in the structures below. This means that there is no obvious structural change that is responsible for variations in rate. This leads to the

conclusion that any changes to the enzymatic rate must have their origins rooted in the dynamic activity of the enzyme.

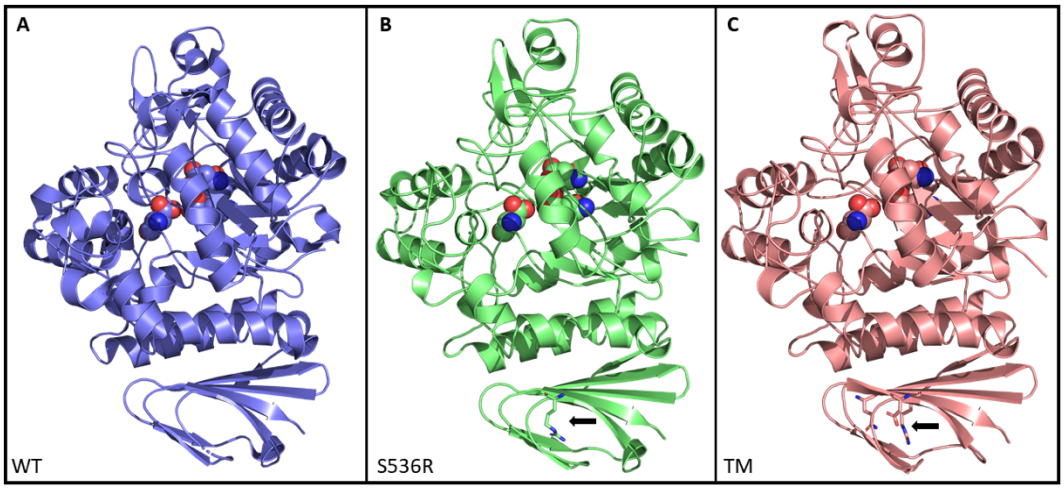

Figure 4.5. Structural comparison of different variants of the MalL enzyme. A) MalL WT (PDB: 4M56). B) MalL S536R. C) MalL TM (S536D, E554Q, V556R). Catalytic residues represented as spheres. Mutations are shown as sticks and indicated with an arrow.

Table 4.5. Root Mean Square Deviation between MalL variants.

	4M56_A	4M56_B	S536R
S536R	0.325 (555)	0.381 (556)	NA
TM	0.570 (555)	0.695 (556)	0.587 (559)

*Values in brackets are number of atoms compared

MalL has three domains which are easily identified in the structures above. The C-terminal domain C is located at the bottom with the catalytic TIM barrel domain A in the centre and the calcium binding domain B at the top.

Some general structural observations between the variants are as follows. In the C-terminal domain, there has been some minor movement in the β-strands which alters the distance between the two β-sheets. In the catalytic domain, some helices shown on the left side have altered their curvature which creates a slightly differently shaped pocket next to the active site. There has been a slight movement of the most distal loop in the calcium binding domain. Calcium binding

is a general feature of many glycoside hydrolase enzymes. The calcium binding in MalL S536R and MalL TM is consistent with the β_1-α_1 loop calcium binding site seen in other GH13 enzymes.

These combinations of minor structural changes in the MalL mutants are proposed to have led to dynamic changes within the enzyme, which has ultimately affected catalysis. This is likely to be due to the changes at the amino acid level that have transmitted structural perturbations from the site of mutation to the site of catalysis.

4.4.1.1 Comparing mutated residues with urea binding

The mutations in MalL S536R and MalL TM were designed to mimic the effects of urea binding in the C-terminal domain of MalL. In both cases an arginine was introduced in a position to occupy the binding site of urea. In the TM mutant two additional mutations were introduced in order to provide support to the introduced arginine. The crystal structure of both mutants reveal that the introduced arginine residue occupies the area that would ordinarily be the binding site for urea (Figure 4.6 and Figure 4.7).

In the single mutant (S536R), the arginine residue protrudes into the space between the β-sheets which is the same space occupied by urea (Figure 4.6). Examination of the binding interaction of urea in the WT structure shows that urea interacts with Ser536 as well as Glu554, which is on the opposite side of the urea molecule. In the S536R mutant, the position of Glu554 is unaltered compared to the wildtype enzyme (Figure 4.6).

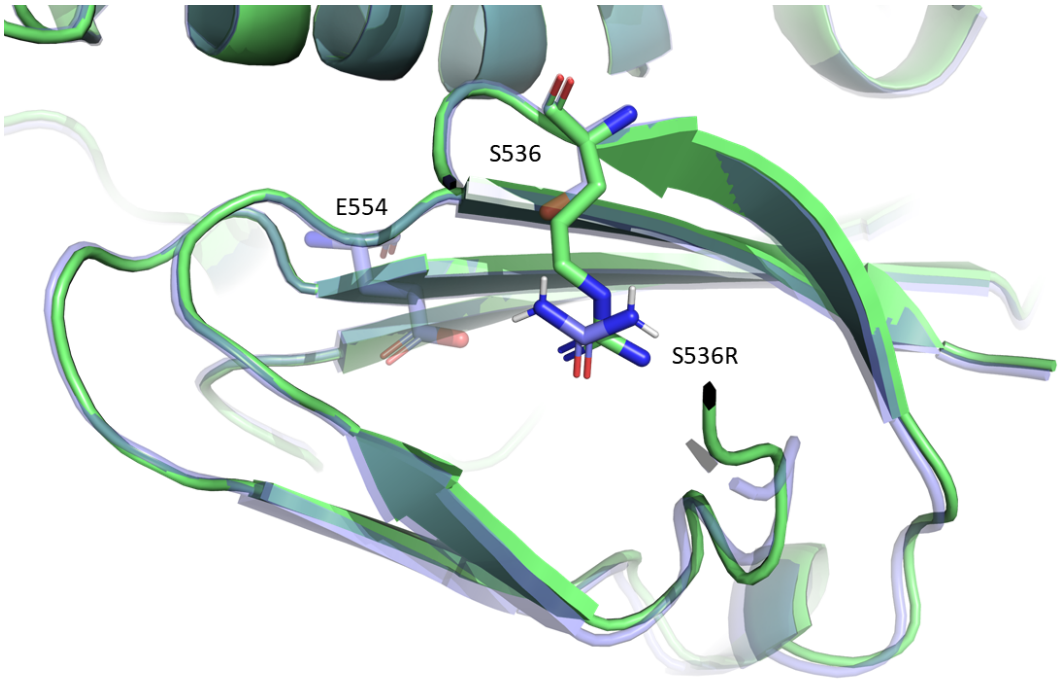

Figure 4.6. Cut-away view of an overlay of MalL S536R with MalL WT with urea bound in the C-terminal domain. MalL S536R is green with Arg536 and Asp554 shown as sticks; MalL WT (with urea bound) is blue with urea and the interacting residues shown as sticks.

In the MalL TM it is also the introduced arginine residue that occupies the same space as the urea molecule (in the WT structure) however the arginine comes from the adjacent β-strand from the V556R mutation (Figure 4.7). The aspartic acid introduced in place of Ser536 does not occupy the urea binding site and the third mutation of E554Q is nearby.

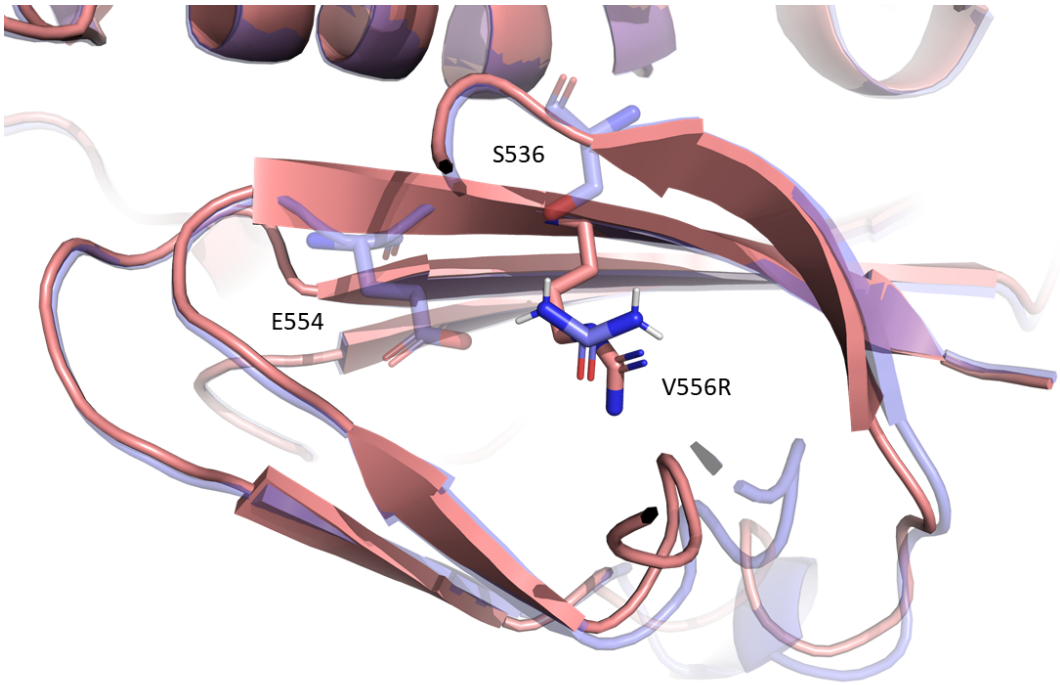

Figure 4.7. Cut-away view of an overlay of MalL TM with MalL WT with urea bound in the C-terminal domain. MalL TM is pink with mutated residues shown as sticks; MalL WT (with urea bound) is blue with urea and the interacting residues shown as sticks.

4.4.2 B-factor analysis of MalL variants

B-factors describe the decrease in scattering intensity due to thermal motion. In general B-factors can be used as an indicator for the relative vibrational motion of atoms contained within a protein, where a low B-factor is indicative of reduced motion in a well-ordered site (Sun *et al.*, 2019). B-factor analysis of MalL variants are shown in Figure 4.8.

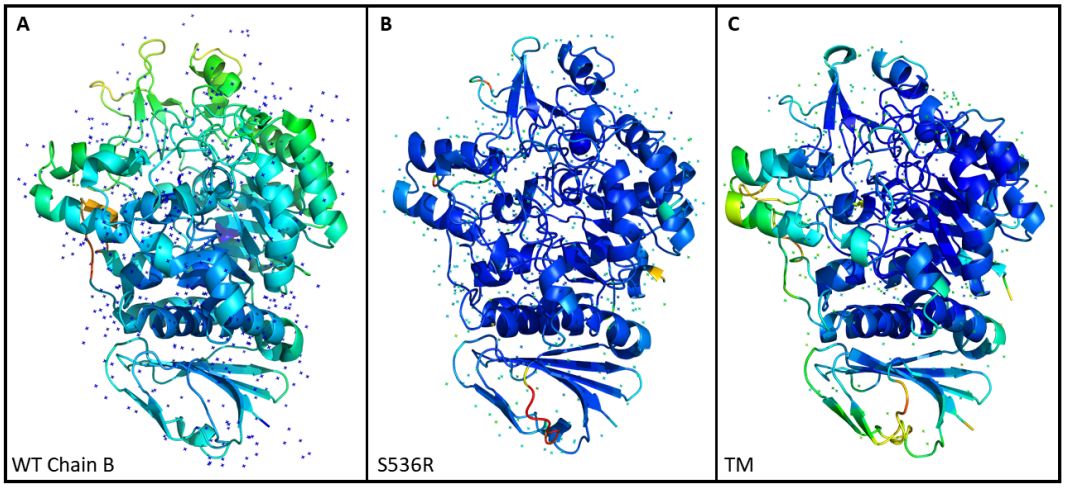

Figure 4.8. Average B-factor analysis of MalL variants viewed by their respective 3D structures. A) MalL WT, 2.3 Å, B-factor range 11.2 – 89.8. B) MalL S536R, 1.10 Å, B-factor range 8.1 – 245.5. C) MalL TM, 1.7 Å, B-factor range 15.2 – 60.6. B-factors shown in scale, from low to high (blue-light blue-green-yellow-orange-red). Colours are indicative of variation within individual structures only.

Overall, S536R shows a significant reduction in B-factors across the entire enzyme compared to WT and TM, except in a few small regions where B-factors are significantly higher, due to limitations in model fitting to regions of poor electron density. This is consistent with a significant rigidification of the enzyme dynamics which is consistent with the observed kinetic data, which showed an increase in $\Delta C_P^{\ddagger}$ and an increase in Topt.

MalL TM in contrast shows a slight rigidification through domain B (residues 108-171) compared to WT. It also shows a loosening in the second half of the enzyme from β-strand 7 which comprises the domain interface between domains A and C, as well as through the C-terminal domain. This region also contains an α-α extension (residues 390-420) between β_8 and α_8 of the TIM barrel, which forms part of the structure for the substrate binding pocket (Figure 4.9).

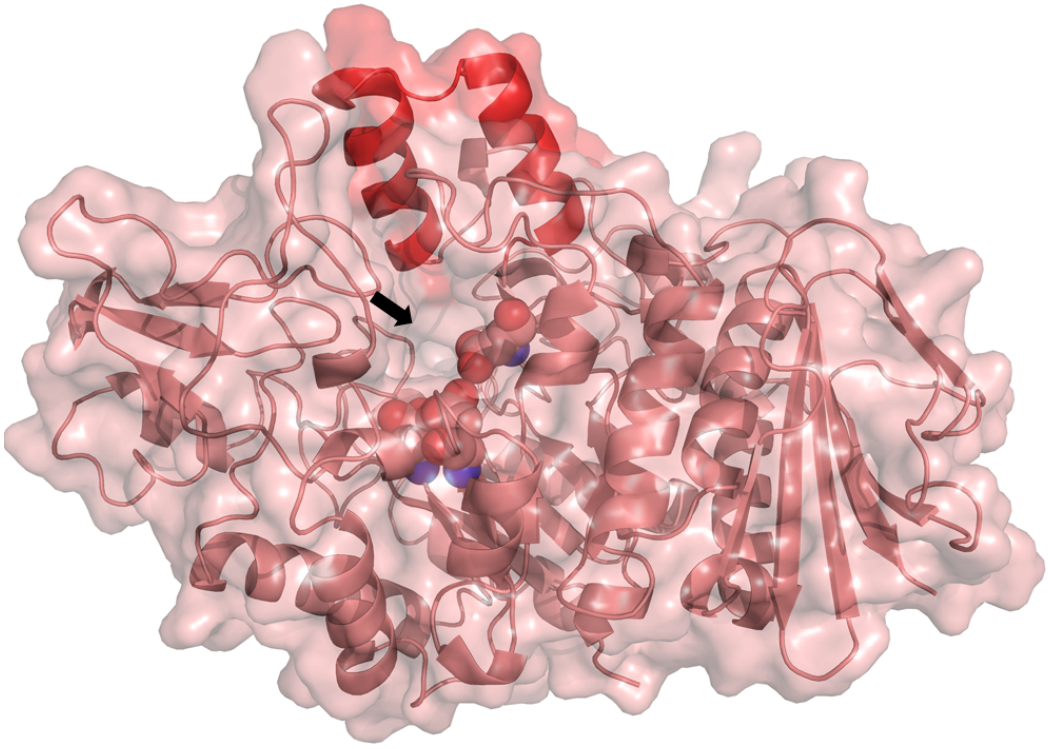

Figure 4.9. Substrate binding pocket of MalL TM. Active site residues are shown as spheres. Mobile α-α extension is coloured red. Binding pocket is between active site residues and α-α extension, indicated by a black arrow.

An alternative way of viewing the B-factor analysis is to graph this variable against the individual amino acids (Figure 4.10). With this view, we can compare the MalL mutants in a numerical fashion instead of by colour. It becomes obvious that MalL S536R has the lowest B-factor throughout the whole protein and the triple mutant has B-factors somewhere in between WT and S536R, depending on the area of the protein. In fact, for the TM the C-terminal region clearly has a higher B-factor than WT, highlighting the dynamic change that has occurred in this β-sheet Greek key domain (section 4.4.3).

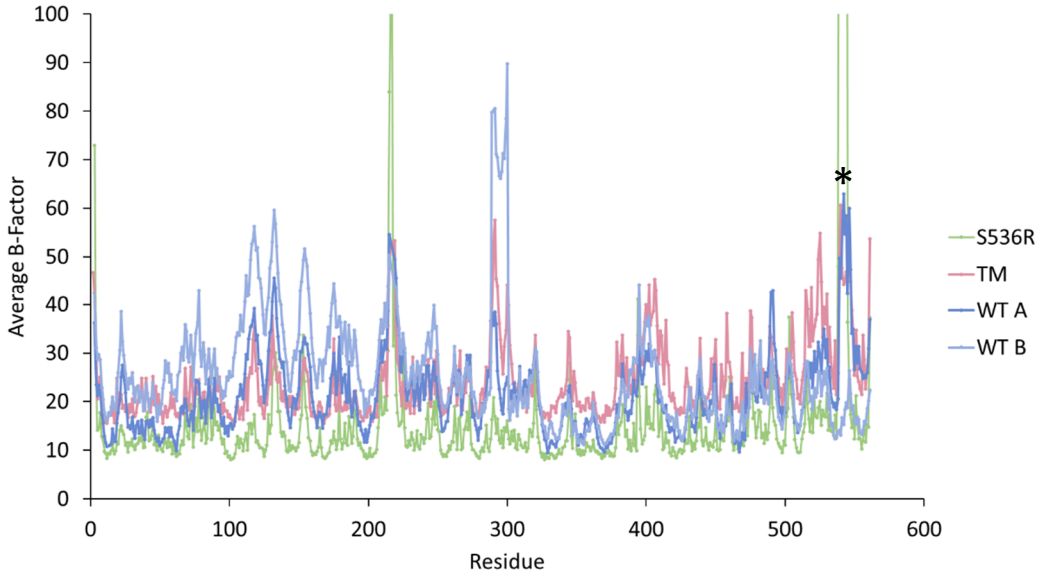

Figure 4.10. Average B-factor analysis of MalL crystal structures shown across the residue range. MalL S536R (Green), MalL TM (Pink), MalL WT Chain A (Dark blue), MalL WT Chain B (Light blue). High B-factors on MalL S536R (residues 215-217 and 539-544) are caused by limited model fit due to poor electron density in highly mobile regions. Asterix represents a specific loop region discussed in section 4.4.3.

4.4.3 Enzyme dynamics of the C-terminal domain

The C-terminal domain of MalL forms a double antiparallel β-sheet in a Greek key motif. This overall structure is consistent between the different MalL mutants. In both S536R and TM the mutated residues are located in this C-terminal domain. Kinetic data from the S536R and TM mutants show that the interactions in this region can have significant effects on the rate and the temperature-dependence of the rate. Both mutants are expected to modify enzyme activity by altering the dynamics of the enzyme. The S536R and TM mutants are hypothesised to exert their effects by rigidifying the C-terminal domain and increasing the flexibility respectively. The B-factor analysis does show an increase in flexibility in the C-terminal domain of the TM and a decrease in the C-terminal domain of S536R (with respect to WT) (Figure 4.11). The interactions of the mutated residues in MalL S536R and MalL TM are shown in Figure 4.12.

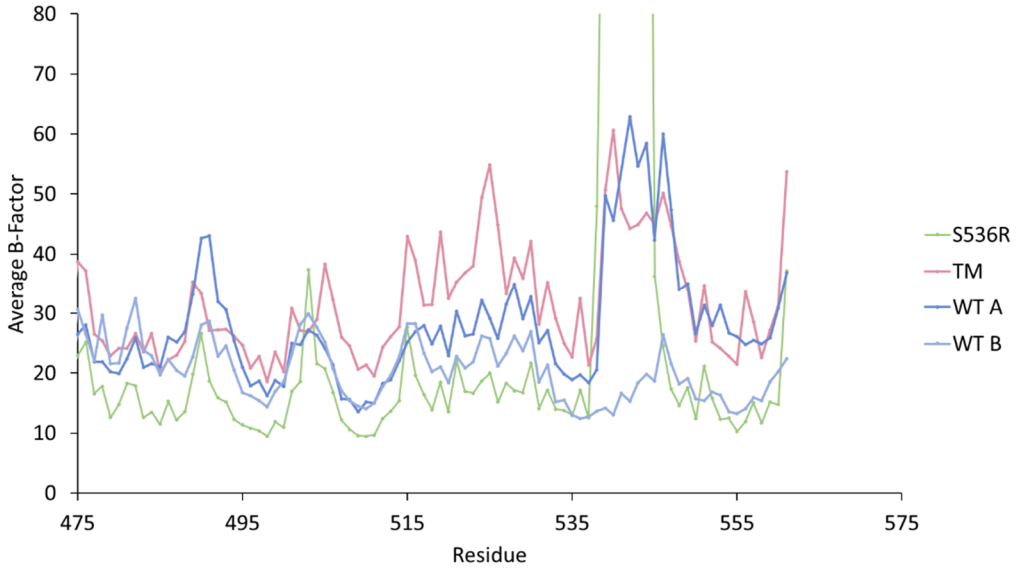

Figure 4.11. Average B-factor analysis of the C-terminal domain of MalL crystal structures. MalL S536R (Green), MalL TM (Pink), MalL WT Chain A (Dark blue), MalL WT Chain B (Light blue).

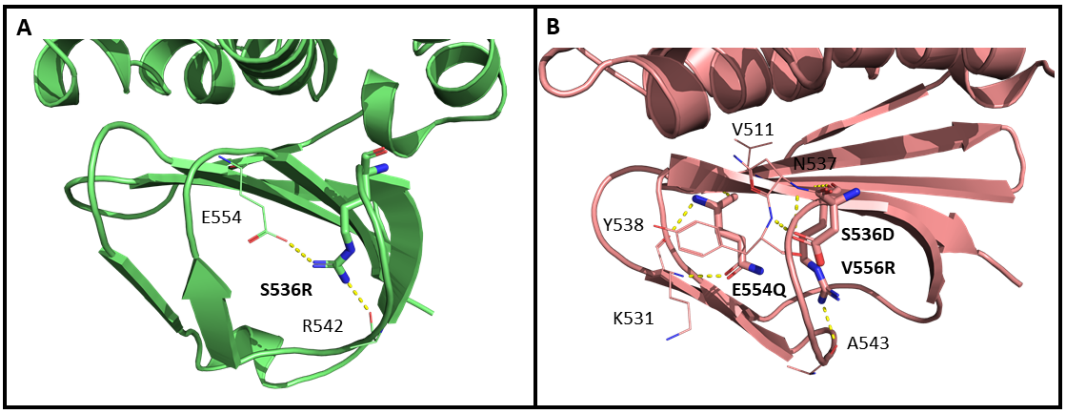

Figure 4.12. Polar interactions of the mutated MalL residues in the C-terminal domain. A) MalL S536R. B) MalL TM. Mutated residues shown as thick sticks and the interacting residues shown as thin sticks with yellow dashes as interactions.

The arginine in the S536R mutant interacts with Glu554 and the backbone of Arg542, creating a connection between a β-strand on one side and a loop on the other side of the Greek key motif. The MalL TM in contrast mainly consists of local interactions. Mutated residue S536D interacts with the adjacent Asn537 and mutated residue E554Q interacts with Lys551 in the same loop and Val511 in the adjacent β-strand. The third mutation (V556R) interacts with the backbone of

Ala543 on the opposite loop in a similar way to the arginine in the S536R mutant. Overall, it appears that the single S536R mutation has rigidified the C-terminal domain by forming a bridge that spans from one side to the other. This is in contrast with the TM where the three mutations have allowed for movement to translate between them.

The specific loop region between Asn537 and Lys546 is generally highly mobile in MalL structures and poorly defined in the electron density, due to its position as a flexible loop on the exterior of the protein. Interestingly in the TM structure this region is well defined. In the TM the E554Q and V556R mutations interact with the S536D mutation and the backbone of Ala543 in the mobile loop to restrict its movement (Figure 4.13). From the illustration of the overlay of the TM with WT MalL it is obvious that these three interacting mutated residues have drawn the mobile loop in towards the centre of the Greek key motif hence contributing to a reduction in movement in this specific area.

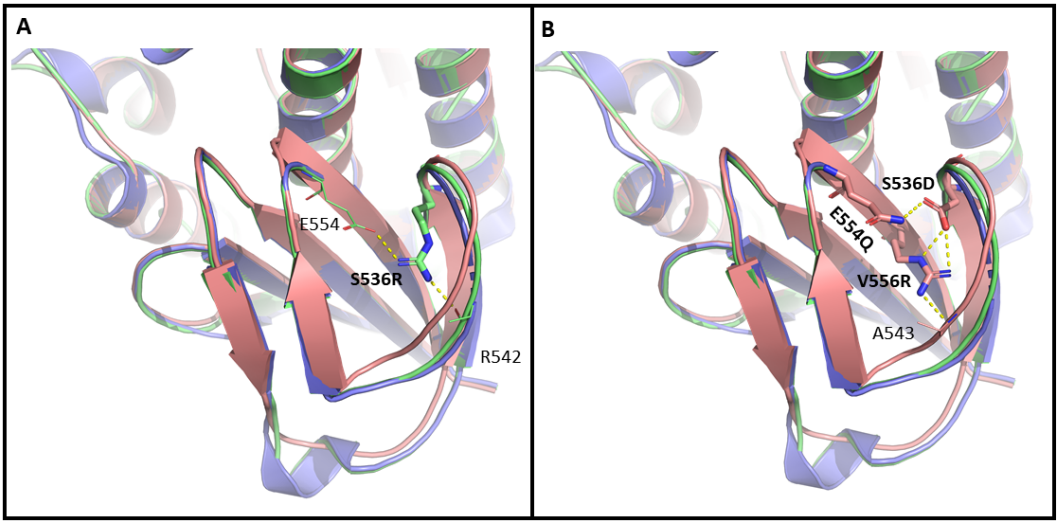

Figure 4.13. Polar interactions of the mutated MalL residues in the mobile C-terminal loop (Asn537-Lys546). A) MalL S536R. B) MalL TM. Mutated residues shown as thick sticks and the interacting residues shown as thin sticks with yellow dashes as interactions. Overlay shows MalL S536R in green, TM in pink and WT in purple.

4.4.4 Structural analysis and MMRT

MalL S536R shows a decrease in flexibility across the whole enzyme consistent with the observed increase in $\Delta C_P^{\ddagger}$. An overall increase in rigidity moves the vibrational modes to higher frequencies and results in a lower heat capacity. This lowers the difference in heat capacity between the reactant state and the more restricted transition state, thus increasing $\Delta C_P^{\ddagger}$. This has also resulted in a lowered enzymatic rate as observed with MalL S536R.

MalL TM shows an overall increase in flexibility compared to the WT enzyme. This is reflected in the kinetic data where TM has a lower $\Delta C_P^{\ddagger}$, and a corresponding increased rate. MalL TM also contains a more flexible α-α loop that forms part of the structure of the substrate binding site. This could possibly be the cause for the observed increase in substrate inhibition. Potential mechanisms for substrate inhibition, as observed in other glycoside hydrolase enzymes include the formation of an Enzyme-Substrate-Substrate complex or a limited turnover Substrate-Intermediate complex (Liao *et al.*, 2016). Both of these mechanisms may be enhanced by increased permeability of substrate to the binding pocket. At low substrate conditions the rate would not be affected as it is not limited by substrate binding.

5 Chemical rescue of MalL activity to mimic allostery

5.1 Introduction

The success of engineering an allosteric mechanism of regulation into previously non-allosteric proteins has to date been limited. To date, successful attempts at engineering have depended on the modularity of protein domains. This approach has utilised naturally occurring allosteric domains and fused them to other separate catalytic domains. However, the success of this approach is limited and requires screening of multiple domain configurations to find successful combinations. Other approaches include designing catalytic activity into an allosteric domain however computational restrictions limit the feasibility of this approach (Deckert *et al.*, 2012).

An alternative strategy (instead of altering an allosteric domain for catalysis), is to alter a catalytic domain for allosteric control by careful design. One approach to achieve this is to design a cavity-forming mutation with the aim of supplementing the enzyme with small hydrophobic ligands to occupy the cavity. If a cavity-forming mutation can be placed in such a way that catalytic activity is affected, then subsequent addition of a small molecule complementary to the formed cavity can restore activity. For full restoration of activity, precise rescue of structure needs to be achieved (Deckert *et al.*, 2012). This specificity can be achieved by muting a scaffolding amino acid such as tryptophan. The cavity can then be rescued by the addition of indole to mimic the tryptophan side chain (Figure 5.1).

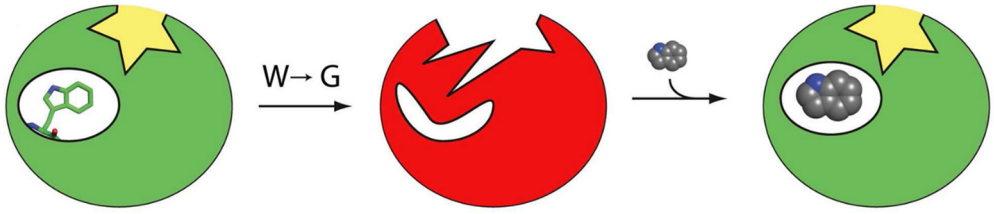

Figure 5.1. Cavity forming mutations (for example tryptophan (W) to glycine (G)) can alter enzyme structure (red) leading to changed enzyme rates. Structural changes can then be rescued by the addition of a chemical mimic of the removed residue.

Deckert and colleagues explored tryptophan mutations as a way to alter active site geometry and thus alter enzyme rate (Deckert *et al.*, 2012). Follow up studies also indicated that similar tryptophan mutations can alter enzyme rates by alternative mechanisms such as enhanced fluctuations (alteration of the conformational ensemble) or localised unfolding (Xia *et al.*, 2013). We therefore attempted to apply this approach with our own isomaltase model enzyme. The isomaltase enzyme (MalL, EC: 3.2.1.10) catalyses the hydrolysis of various disaccharides (such as maltose or sucrose) and a range of maltodextrins. MalL consists of two domains, an N-terminal catalytic TIM (triosephosphate isomerase) barrel domain and a C-terminal all β-strand domain. We intended to replicate the chemical rescue technique described above in MalL. A buried tryptophan residue was identified at position 531 in the C-terminal β-strand domain for mutagenesis. The C-terminal domain has previously been shown in our earlier research to have effects on catalytic rate presumably due to dynamic effects (Section 3) and hence was an obvious target for mutagenesis. This residue was mutated to glycine and alanine and characterised kinetically with and without the addition of indole as a chemical rescue agent. The position of the mutated tryptophan residue 531 is shown in Figure 5.2. Interactions of Trp531 are shown in Figure 5.3.

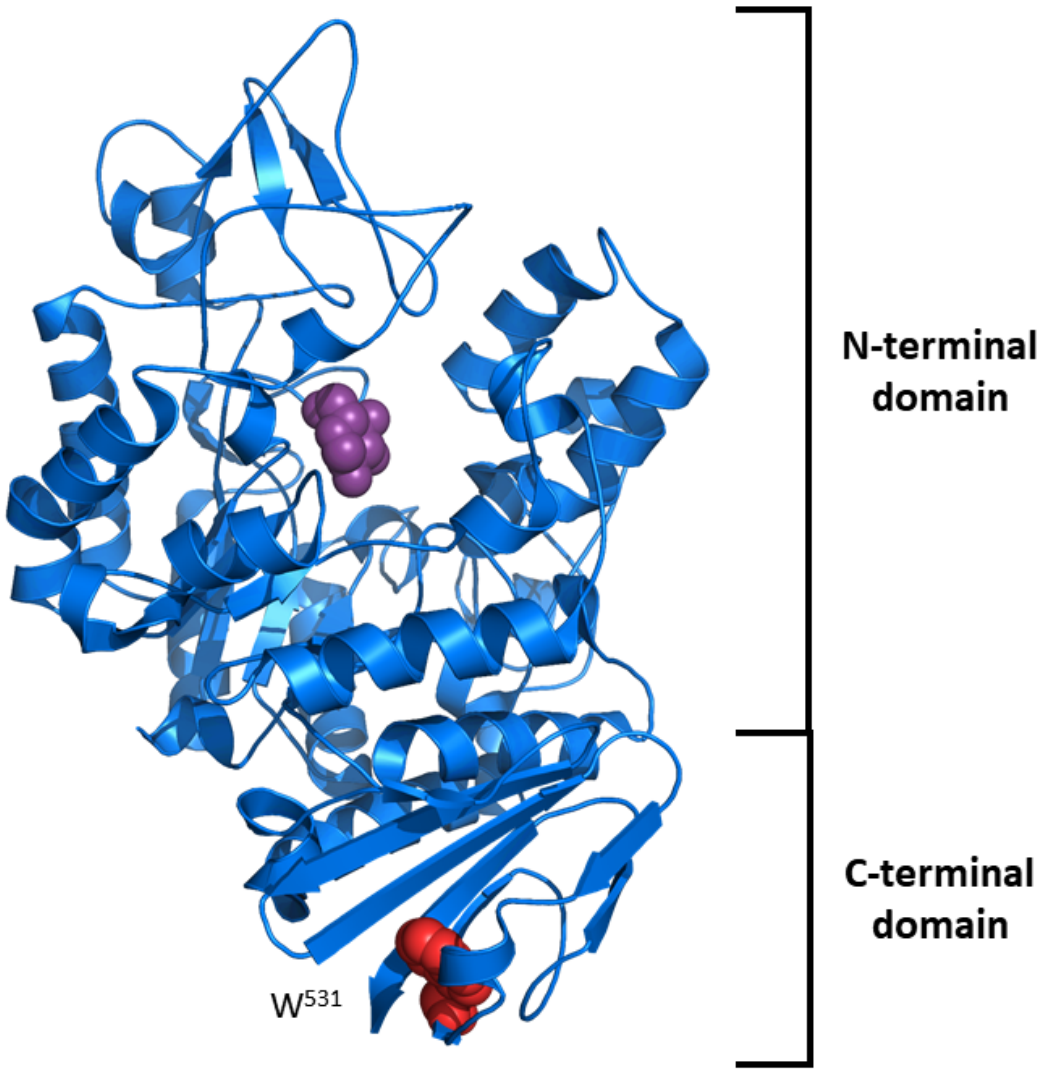

Figure 5.2. Position of the W531G mutation in wildtype isomaltase enzyme MalL [PDB 5WCZ (van der Kamp et al., 2018)]. The mutated tryptophan (W^{531}) is shown as red spheres in the C-terminal domain. Substrate analogue (1-Deoxynojirimycin) is shown as purple spheres in the N-terminal domain.

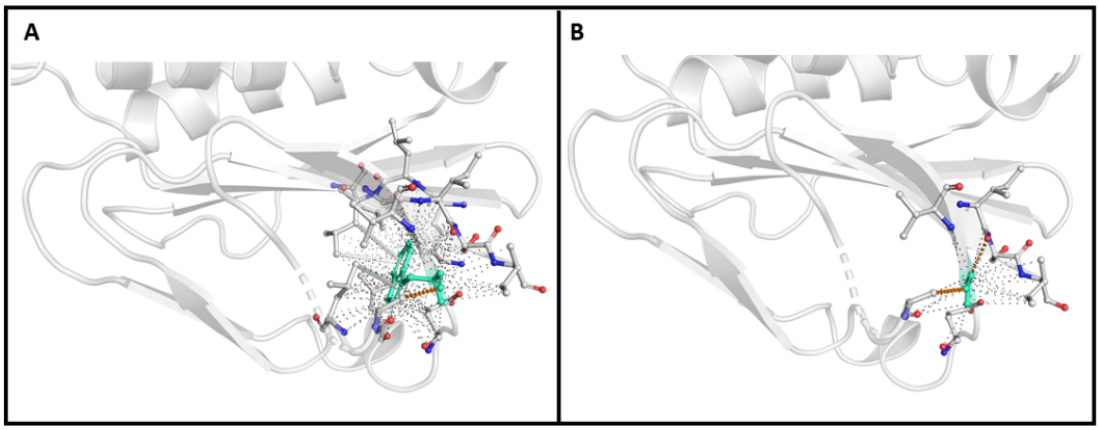

Figure 5.3. Interactions of W531 and G531 in the C-terminal domain of MalL. A) MalL W531 (Blue). B) MalL G531 (Blue)

5.2 Mutagenesis and cloning of MalL tryptophan mutants

5.2.1 Two-halves PCR

Mutation of tryptophan 531 to either alanine or glycine was achieved using two-halves PCR mutagenesis (Section 2.3.1.2). In this method, initially two PCR reactions are performed which amplify the gene in two sections, introducing the desired mutation in an overlapping region of the two products. A second PCR reaction (using the product of the two initial PCR reactions as a template) allows the mutated halves to anneal and form a complete mutated gene. Mutation of tryptophan 531 generates a forward product 1714 bp long and a reverse product 171 bp long, with the complete joined sequence being 1852 bp long (Figure 5.4).

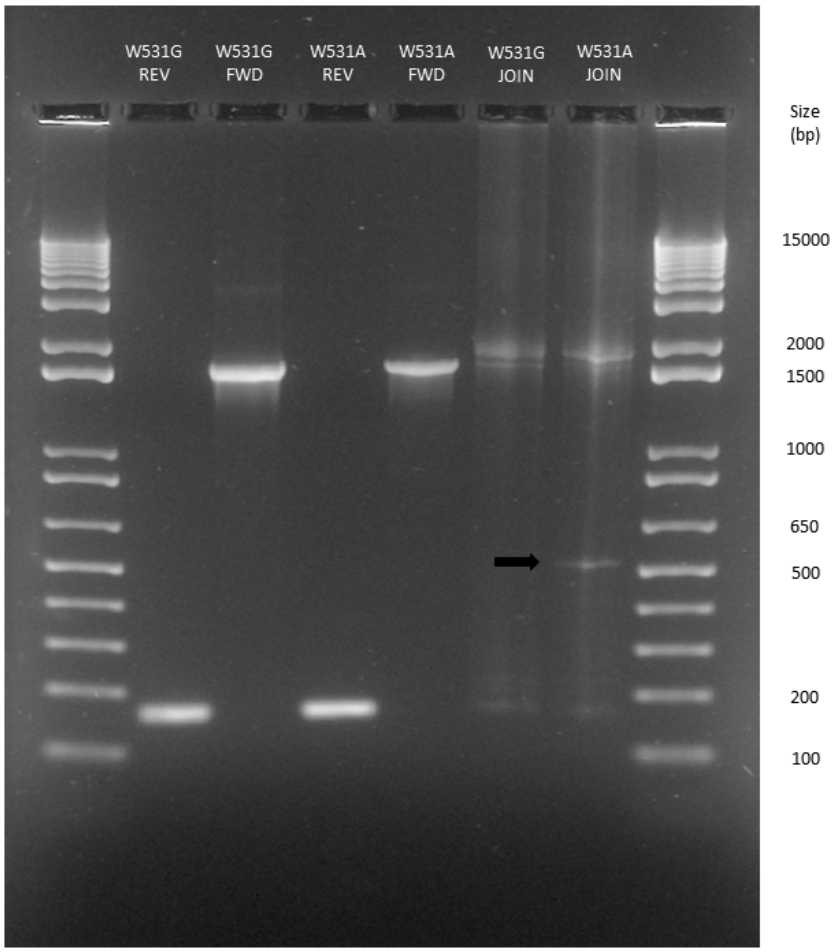

Figure 5.4. PCR products of W531G and W531A forward, reverse and joining reactions on 2% TAE agarose gel.

Bands of the expected sizes (1714, 171 bp) for the forward and reverse reactions could be identified on the agarose gel. The PCR joining reaction however showed a major band at the expected size (1852 bp), as well as additional bands representing the forward and reverse products that were used as the template and other nonspecific bands (Figure 5.4, arrow as an example). Because of this result, full length PCR products were separated by agarose gel electrophoresis (Section 2.3.2) and the appropriate sized band was excised and purified to isolate the complete mutated gene. The mutated gene was then ligated (Section 2.3.4) into the pPROEX-Htb expression vector and transformed into the *Escherichia coli* strain DH5α (Section 2.3.5). Transformed colonies were selected by growth on LB AMP plates and screened for the 1852 bp gene insert by colony PCR (Figure 5.5).

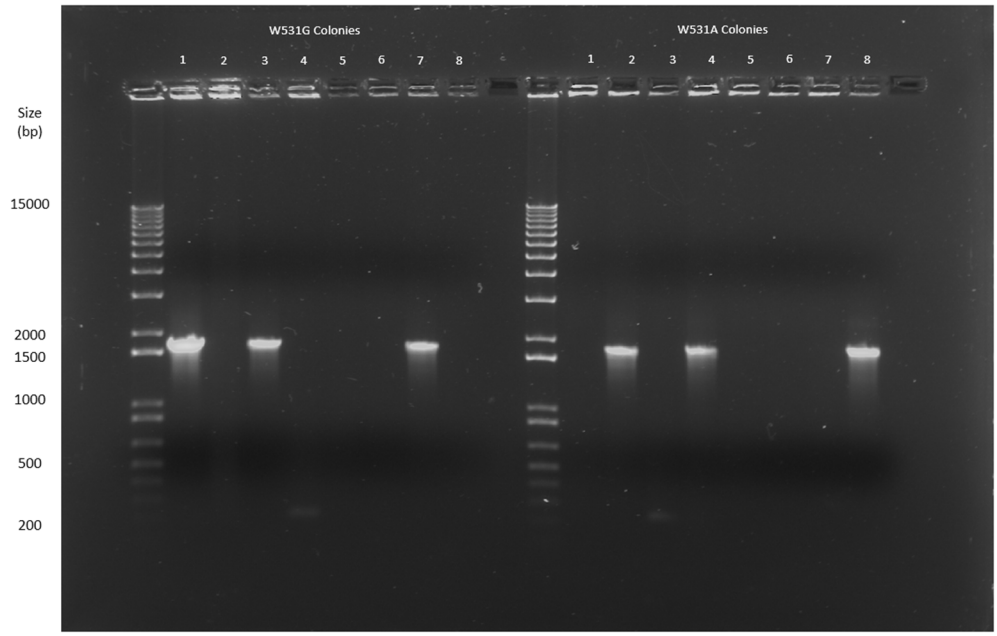

Figure 5.5. Agarose gel products of a colony PCR of the gene for MalL mutants. Eight colonies were selected from each of the transformations of W531A and W531G into E. Coli. Products were run on a 1% TAE agarose gel.

Three colonies from each mutation (W531A or W531G) had the expected gene insert represented by the 1852 bp band. Colonies containing the gene insert were sent for sequencing. Only one sequence for W531G had sufficient quality or lacked other mutations to confirm the inserted mutation. Mutation to alanine proved problematic as additional mutations were introduced during PCR and therefore it was unable to be characterised kinetically. The successful sequencing of MalL W531G is shown in Figure 5.6.

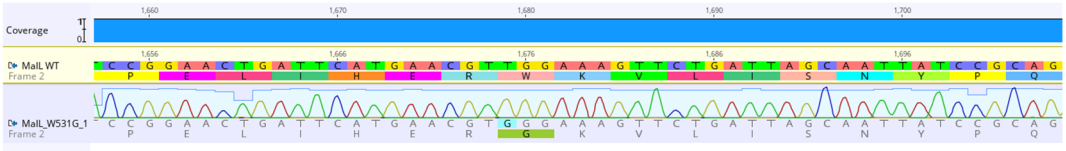

Figure 5.6. Sequencing result showing successful W531G mutation in MalL.

5.3 Expression trials of MalL W531G

MalL W531G was expressed and purified using well established protocols for MalL variants (Section 2.1). MalL W531G was expressed in *E. coli* in the pPROEX-Htb expression vector under the control of the *lac* promoter. Protein expression was induced by addition of IPTG. Soluble protein was obtained by purification by immobilised metal ion affinity chromatography (IMAC) and purity confirmed by polyacrylamide gel electrophoresis (PAGE) (Figure 5.7).

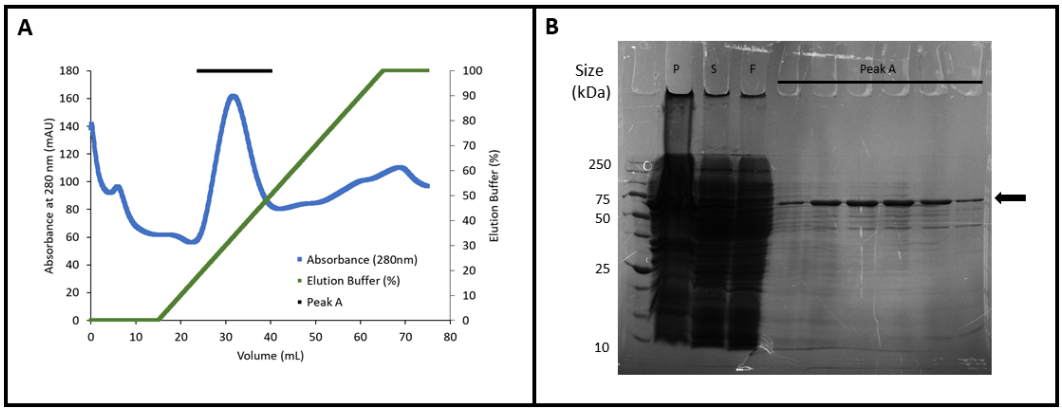

Figure 5.7. Ni IMAC Purification of MalL W531G A) Chromatogram of MalL W531G Ni IMAC purification. B) PAGE gel of MalL W531G Ni IMAC purification. P = Insoluble pellet, S = Supernatant containing protein loaded onto column, F = Flow-through of column minus bound target protein. Arrow indicates intense band showing purification of target protein.

MalL W531G protein eluted from the IMAC column as a single peak at approximately 20% elution buffer. This peak contained a mixed population of protein with MalL comprising approximately 90% of the total protein as identified by SDS-PAGE. An intense band between 50 and 75 kDa (represented by a black arrow) corresponded to the expected MalL size of 69.5 kDa. MalL W531G was further purified by size exclusion chromatography (SEC) and purity confirmed by PAGE. MalL W531G eluted from the SEC column with improved purity between 65 and 80 mL which is estimated to be a molecular weight of approximately 69.5 kDa indicating MalL purified as a soluble monomer (Figure 5.8). A band between 50 and 75 kDa can be seen on the gel corresponding to the expected size of MalL.

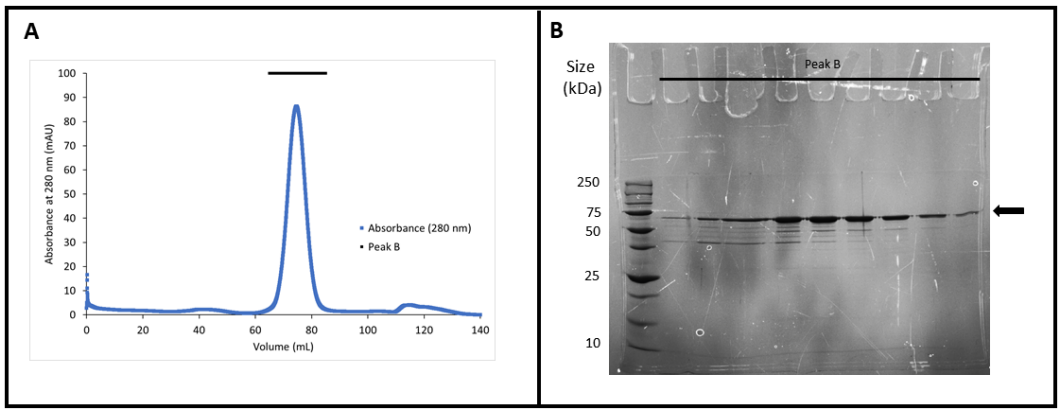

Figure 5.8. PAGE Gels of MalL W531G purification. A) Ni IMAC of MalL W531G. Samples A5-13 correspond to a range of fractions over peak B) SEC of MalL W531G. Samples B1-9 correspond to fractions over peak. Arrow indicates intense band showing purification of target protein.

5.4 Characterisation of MalL W531G

A range of kinetic assays were performed based on established protocols for MalL variants (Section 2.2) using the substrate *p*-nitrophenyl-α-D-glucopyranoside (PNG).

5.4.1 Michaelis-Menten Kinetics

The Michaelis-Menten kinetic model was used to find the k_{cat}, K_M and K_i for MalL W531G with and without the addition of indole. Two indole concentrations were chosen based on previous work by Deckert *et al*., (2012). Deckert and colleagues found that addition of indole at 10 mM showed complete catalytic rescue of the enzyme structure (Deckert *et al.*, 2012). Indole concentrations below this value at 5 mM and 0.5 mM were therefore chosen to assay indole effect on MalL W531G (Figure 5.9).

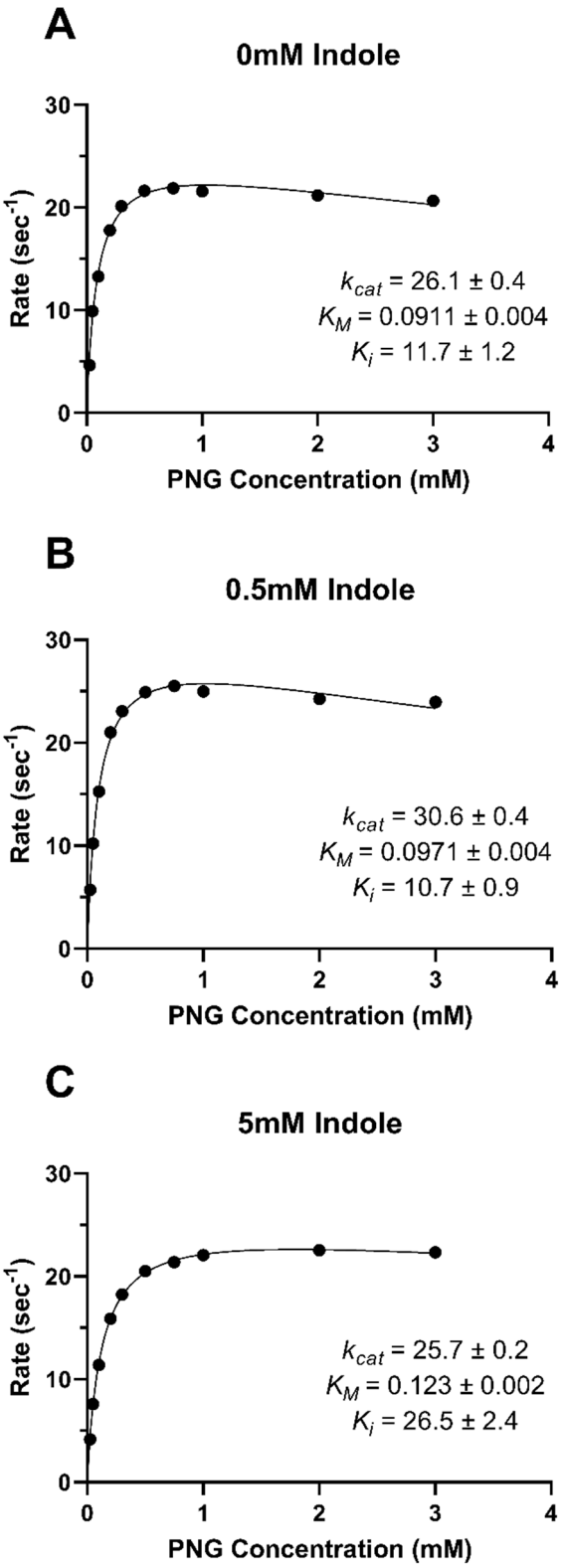

Figure 5.9. Michaelis-Menten kinetics of MalL W531G at 25°C with varying Indole effector concentrations. A) 0 mM indole. B) 0.5 mM indole. C) 5 mM indole. Data fit to substrate inhibition model. PNG is MalL substrate (para-nitrophenyl-α-D-glucopyranoside). Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

The k_{cat} was significantly increased at low indole concentrations (0.5 mM) but decreased slightly at high concentrations of indole (5 mM). K_M increased with indole concentration. At 5 mM indole K_M was nearly 35% higher than at 0 mM indole. At 0 mM and 0.5 mM indole MalL W531G showed low levels of substrate inhibition with K_i's of 11.7 ± 1.2 mM and 10.7 ± 0.9 mM respectively. Substrate inhibition was observed to a lesser degree at 5 mM indole, with a K_i of 26.5 ± 2.4 mM. A non-substrate inhibition model was considered; however, the data was better fit by a substrate inhibition model (Figure 5.10).

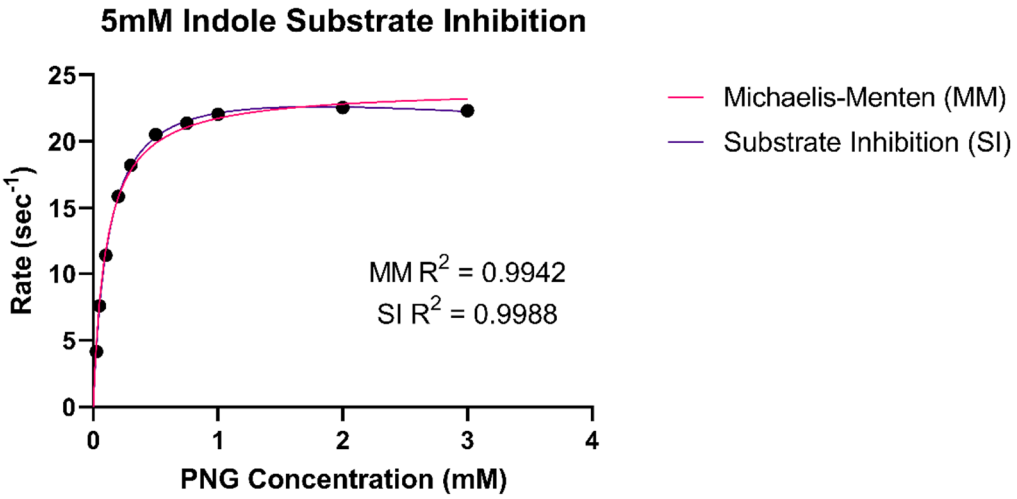

Figure 5.10. MalL W531G 5mM Indole Substrate Inhibition Michaelis-Menten fit

Table 5.1. Michaelis-Menten constants for MalL W531G Michaelis-Menten kinetics with varying indole concentrations.

	Indole Concentration		
Parameter	**0 mM**	**0.5 mM**	**5 mM**
k_{cat} (sec^{-1})	26.1 ± 0.4	30.6 ± 0.4	25.7 ± 0.2
K_M (mM)	0.0911 ± 0.004	0.0971 ± 0.004	0.123 ± 0.002
K_i (mM)	11.7 ± 1.2	10.7 ± 0.9	26.5 ± 2.4

5.4.2 Determination of Indole Inhibition Concentration

Given the improvement of MalL activity seen with 0.5 mM indole (Section 5.4.1) we wanted to find the concentration of indole that would give optimal MalL activity. The activity of MalL W531G was therefore assayed in 0.4 mM increments up to 4.0 mM indole concentration using a PNG substrate concentration of 1mM (approximately 10 X K_M) (Figure 5.11).

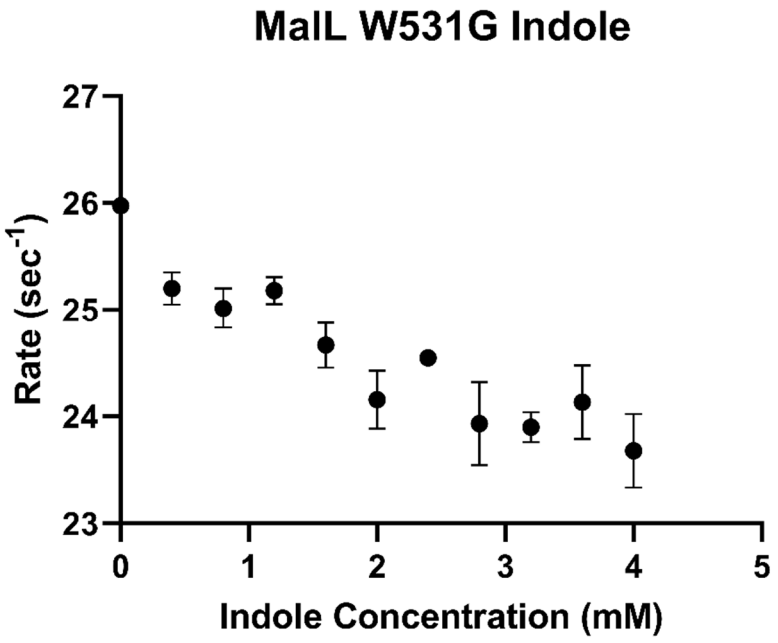

Figure 5.11. Increasing the concentration of indole in a standard assay inhibits the activity of MalL W531G at 25°C. Error bars were visible are the standard deviation of three replicates.

MalL W531G was found to be inhibited by Indole at as low as 0.4 mM in concentration. This is in contrast to the data presented above (Section 5.4.1), where indole was found to act as an activator at a concentration of 0.5 mM. Activity is progressively decreased with increasing concentrations of indole. These data are preliminary and further work is required to resolve the differences between these experiments.

5.4.3 Temperature Characterisation of W531G

The response of MalL W531G to temperature was measured as described in Section 2.2.2. The data was fitted with a temperature dependant $\Delta C_P^{\ddagger}$ equation.

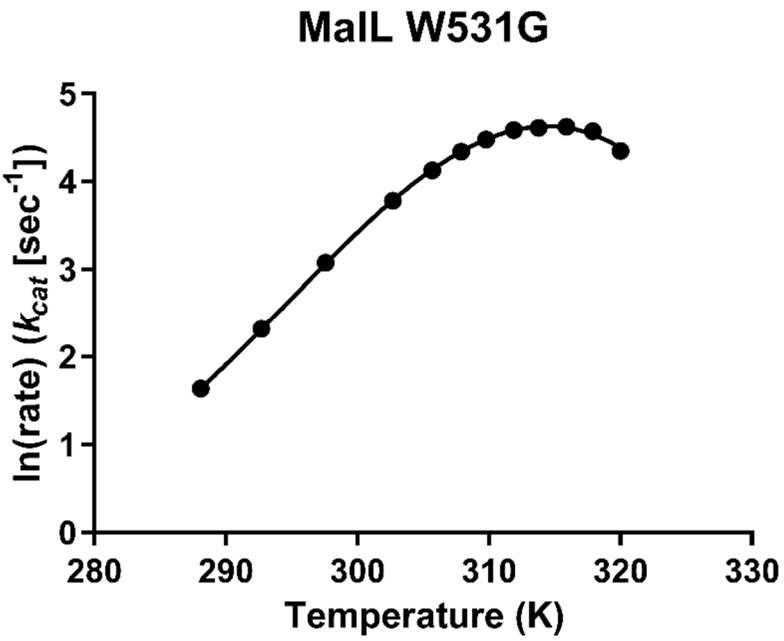

Figure 5.12. Temperature characterisation of MalL W531G fit with the MMRT equation using a temperature dependent $\Delta C_P^{\ddagger}$. Points are averages of triplicate values. Error bars where visible are the standard deviation of three replicates.

MalL W531G shows a curved temperature dependence consistent with MMRT (R^2 = 0.9976). MalL W531G has a temperature optimum of 314.6 K (41.6°C) and the rate is temperature dependant with a $\Delta C_P^{\ddagger}$ of -9.6 ± 0.1 kJ mol^{-1} K^{-1} (T_0 = 310.6 K) (Table 5.2). Compared to MalL wildtype MalL W531G has a larger $\Delta C_P^{\ddagger}$ with wildtype at -8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K) (Figure 5.13).

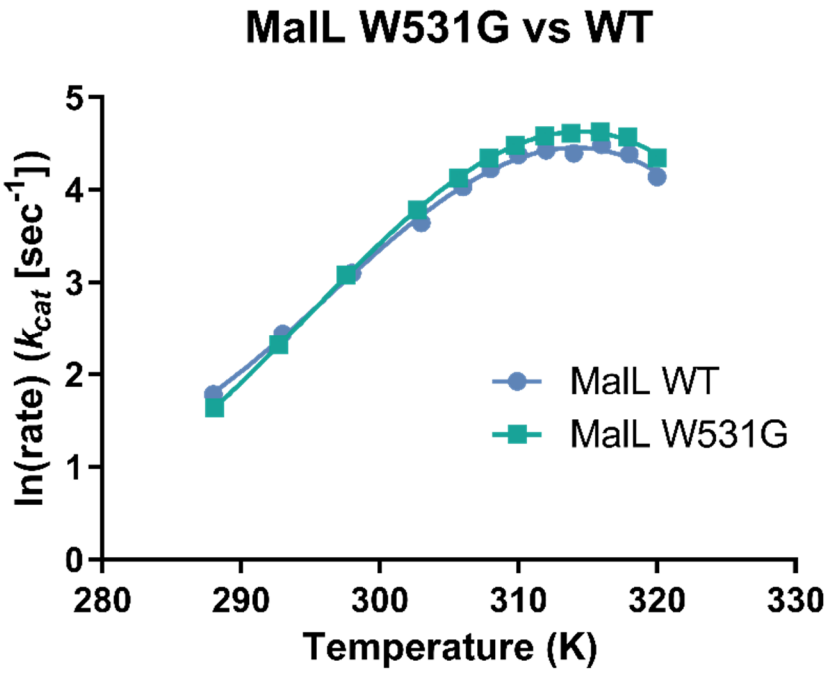

Figure 5.13. MalL comparison between W531G and Wildtype

Table 5.2. MMRT fit parameters of MalL W531G and MalL Wildtype

	Enzyme	
Parameter	**MalL WT**	**MalL W531G**
T_0 (K)	309.4	310.6
$\Delta H^{\ddagger}$ (T_0) (kJ mol^{-1})	42.9 ± 1.2	39.3 ± 0.5
$\Delta S^{\ddagger}$ (T_0) (J mol^{-1} K^{-1})	-70.7 ± 3.7	-81.4 ± 1.5
$\Delta C_P^{\ddagger}$ (T_0) (kJ mol^{-1} K^{-1})	-8.5 ± 0.3	-9.6 ± 0.1
A (kJ mol^{-1} K^{-2})	-218 ± 15	-221 ± 6
T_{opt} (K)	313.4	314.6

5.5 Discussion

A tryptophan to glycine mutation was able to be successfully mutated in MalL. Although this mutation does have an effect on the rate, the full extent of the interactions with indole was unable to be fully characterised. Overall W531G performed reasonably similarly to the wildtype enzyme in the absence of indole.

It was expected that the removal of tryptophan would alter the dynamics of the C-terminal domain. Given the results of the S536R mutant where we saw a decrease in the rate that could be attributed to a rigidified C-terminal domain, we thought that a tryptophan mutant could see a change in the enzymatic rate.

The tryptophan mutants as described by Deckert and colleagues functioned by altering the conformation of the active site. Our mutants in contrast aimed to work by altering the dynamics of an accessory domain. Changes to the enzymatic rate were confirmed by Michaelis-Menten assays which showed that the addition of 0.5 mM indole increased k_{cat}. Changes were also observed in the K_M and K_i, which indicates a complex dynamic effect is changing the enzyme rate.

MalL W531G had a larger $\Delta C_P^{\ddagger}$ than MalL wildtype indicating that the dynamics of the reactant state versus the transition state in MalL W531G has been altered. The negative $\Delta C_P^{\ddagger}$ in enzymatic reactions comes from a rigidification of the enzyme's dynamics in the transition state. The cause of this increase in $\Delta C_P^{\ddagger}$ could come from an increase in the dynamics of the C-terminal domain in the reactant state. Addition of indole would therefore restore rigidity to this domain and decrease $\Delta C_P^{\ddagger}$ which would be consistent with the results in the indole concentration assay (Section 5.4.2). Normal mode analysis (Section 2.5) predicts an overall increase in flexibility in the C-terminal domain (Figure 5.14) (Rodrigues *et al.*, 2018).

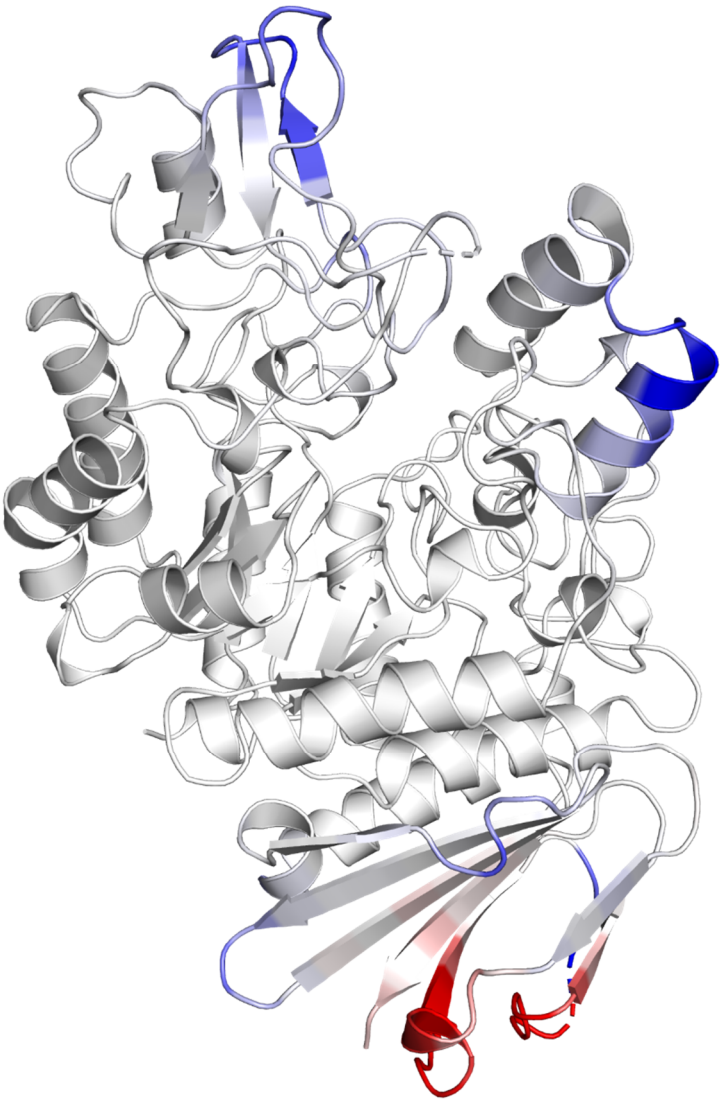

Figure 5.14. Predicted dynamic changes of MalL W531G using normal mode analysis. Regions of increased flexibility are coloured red. Regions of decreased flexibility are coloured blue.

The indole concentration optimisation assay gave contrasting results to the Michaelis-Menten assays. Michaelis-Menten seems to indicate that enzyme activity increases with indole concentration, shown by the increase in k_{cat} between 0 and 0.5 mM indole. This increase appears to be offset by increases in K_M which results in an inhibitory effect, returning MalL W531G to a basal rate at increased indole concentrations (5 mM).

The indole concentration assay in contrast showed small decreases in enzyme activity with increased indole concentration. The decrease in rate over the 4 mM range showed a variation of only 7% (26.5 to 23.7). While the trend is somewhat inconsistent it does show a clear decrease. In addition, given the overall small change in rate over the tested range, error introduces substantial variability that may influence interpretation. Given the contradictory results it is required that these assays be repeated to allow for firm conclusions to be drawn.

Future work would include repeating assays for MalL W531G. In addition, the cloning work for MalL W531A should be repeated. Time limitations meant that that this was not initially able to be completed. Given the relatively low success rate (1 successful mutant from 16 in colony PCR) of mutation a modified PCR protocol with a higher fidelity DNA polymerase may be necessary to prevent additional mutations from being inserted. MalL W531A would allow for characterisation of different binding capacities for indole by reducing the binding complementarity to the binding pocket.

The crystal structure of MalL W531G could provide insights into the mechanism by which indole affects the rate. Differences between the apo and indole bound structures would allow for the effect of indole to be characterised. In addition, the crystal structure will allow for study using molecular dynamics which allows for dynamic effects across the whole protein to be measured *in silico*.

6 Discussion

Enzymes are extraordinary biological catalysts that can accelerate the rate of chemical reactions by incredible amounts. One important aspect of enzyme activity is their response to temperature. Enzymes are very sensitive to temperature changes and produce a distinctive curved rate vs temperature profile. As temperatures increase enzymatic rate also increases, until a point where the rate is at a maximum (called the T_{opt}) followed by a decrease in enzymatic rate. The activity of enzymes with regards to temperature has been traditionally modelled using the Arrhenius and Eyring equations which predict an exponential increase in rate with temperature. Deviations from these models has previously been attributed to enzyme denaturation at elevated temperatures.

It has recently been shown by an increasing body of evidence that these deviations are not able to be sufficiently described by enzyme denaturation (Arcus *et al.*, 2016; Hobbs *et al.*, 2013). Hence, enzymes have a more complicated temperature dependence. Macromolecular rate theory (MMRT) describes the temperature dependence of enzymes by adding a term for the activation heat capacity to the rate equation (Hobbs *et al.*, 2013). The heat capacity of an enzyme is dependent on the distribution of vibrational modes for the enzyme molecule (Gómez *et al.*, 1995). In this study we have used the MMRT equation to fit our enzyme rate vs temperature data because of the way it more accurately models the curved temperature profile. In using the MMRT equation we also have a tool to calculate the change in heat capacity for an enzyme catalysed reaction based on experimental data.

Experimentally, there are a number of approaches we can take to investigate enzyme dynamics, one of which includes the binding of ligands. Ligand binding can influence the vibrational modes of macromolecules, which will alter the enzyme dynamics and therefore alter the heat capacity. Use of MMRT in this study was pivotal in order to determine heat capacity changes of the enzyme and any changes that may occur to it. MMRT has found a wide range of applicability towards biological phenomena including temperature adaptation, and

temperature responses from the enzyme level up to an ecosystem scale (Arcus *et al.*, 2016; Hobbs *et al.*, 2013; Schipper *et al.*, 2014).

Allosteric regulation is a widespread phenomenon observed in many enzymes whereby an enzyme is regulated by an effector binding at a site other than the active site. Despite the widespread use of this type of regulation in biology, the underlying mechanisms of allostery are still poorly understood. We hypothesise that allosteric enzyme regulation can be achieved using a mechanism that alters the heat capacity of an enzyme. To test this, we have developed several mutants of the *Bacillus subtilis* isomaltase enzyme, MalL. MalL has been widely used as a model for MMRT and a number of mutants have already been characterised both kinetically and structurally (Hobbs *et al.*, 2013). In addition to these mutants, MalL has also been characterised bound with urea (unpublished) and bound to a transition state analogue (PDB: 5WCZ) (van der Kamp *et al.*, 2018).

The wild type (WT) MalL enzyme had previously been characterised with low concentrations of urea. This showed that urea could bind to and potentially affect the dynamics of MalL. A urea binding site was identified in the C-terminal domain of MalL far away from the active site. Analysis of the urea binding site by crystallography revealed that bound urea interacted with S536 and E554. Two mutants were developed in collaboration with the Institute for Protein Innovation to mimic the effect of urea binding. In these mutants, either S536 or V556 were mutated to arginine (because of their proximity to the bound urea) to provide atoms that would mimic the urea binding site. Additionally, the V556R mutation was supported by two additional mutations (E554Q, S536D) which together we have termed a triple mutant (TM).

These two mutants were characterised both kinetically and structurally. MalL S536R showed a slight decrease in the rate when characterised using Michaelis-Menten kinetics at 25°C. MalL S536R had a k_{cat} of 25.4 ± 0.5 sec^{-1} compared to WT with a k_{cat} of 27.1 ± 0.6 sec^{-1}. MalL S536R showed a larger decrease in the rate when characterised in a temperature assay. MalL S536R had significantly lower activity than WT across a range of temperatures. MalL S536R had a k_{cat} of 43.2 ± 1.0 sec^{-1} compared to WT of 82.9 ± 1.1 sec^{-1} at 312 K (39°C).

The activity assay of MalL S536R with temperature was well described by MMRT (R^2 = 0.9979) with a slight increase in the T_{opt}. MalL S536R had an increased $\Delta C_P^{\ddagger}$ of -7.4 ± 0.2 kJ mol^{-1} K^{-1} (T_0 = 312.8 K) compared to WT (-8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K)). In addition, $\Delta C_P^{\ddagger}$ was not as strongly dependent on temperature as in the WT enzyme, with an A value of -139 ± 9 kJ mol^{-1} K^{-2} compared to -218 ± 15 kJ mol^{-1} K^{-2} in WT. These results are consistent with an overall rigidification of the enzyme when compared to WT.

The kinetic results were corroborated by analysis of the crystal structure of MalL S536R. MalL S536R showed the same general domain architecture to MalL WT. The S536R mutation successfully occupied the same space as urea (as in the WT urea structure). B-factor analysis suggests an overall increase in rigidity across the entire enzyme. In addition, the crystals diffracted to very high resolution (1.10 Å) and this also suggests that the enzyme molecules are more ordered in the crystal.

Similar to MalL S536R, the MalL TM introduced an arginine at position 556 that also occupied the same space as urea. In contrast to MalL S536R, the TM showed an overall increase in the rate compared to WT. The k_{cat} in the TM was increased to 34.3 ± 1.0 sec^{-1} from 27.1 ± 0.6 sec^{-1} in WT. Similarly, the TM showed an increase in rate in the temperature assay. MalL TM saw a decrease in $\Delta C_P^{\ddagger}$ to -9.3 ± 0.1 kJ mol^{-1} K^{-1} (T_0 = 309.1 K) from -8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K). This is consistent with an increase in flexibility across the enzyme. This was also represented in the crystal structure which showed an increase in flexibility across the second half of the enzyme including the C-terminal domain (as indicated by B-factor values).

In addition to the S536R and TM mutants a third W531G mutant was constructed. Through MalL W531G, we attempted to engineer an allosteric rescue site into the C-terminal domain. Tryptophan is bulky and thus acts as a hydrophobic supporting residue and its removal therefore potentially disrupts the structure and dynamics of the C-terminal domain. Chemical mimicry of the removed tryptophan moiety (by addition of indole) is hypothesised to restore the disrupted structure. Results from the S536R and TM mutants showed that small changes to this domain could have significant effects on the rate.

Kinetic assays of MalL W531G showed conflicting results. In Michaelis-Menten assays with increasing indole concentration, MalL W531G saw activation at low concentrations of indole. In contrast to this, in an indole concentration assay W531G showed increasing inhibition with indole concentration. W531G has a decreased $\Delta C_P^{\ddagger}$ [-9.6 ± 0.1 kJ mol^{-1} K^{-1} (T_0 = 310.6 K)] compared to WT [-8.5 ± 0.3 kJ mol^{-1} K^{-1} (T_0 = 309.4 K)]. This would seem to indicate an increase in flexibility caused by the W531G mutation which may be restored with the addition of indole.

All three mutations were approximately 30 Å from the active site in the C-terminal domain. These mutations had significant effects on the enzymatic rate across this distance. Overall the structures of MalL TM and S536R were very similar to WT and only showed minor deviations in structure. Analysis of RMSD values of MalL TM and S536R to WT chain A (PDB: 4M56) showed values of 0.570 Å and 0.325 Å, respectively. This indicates an overall high degree of similarity between the different structures. Changes to the rate can therefore be attributed to changes in enzyme dynamics and the enzymes heat capacity.

This study hypothesised that allosteric enzyme regulation can be achieved using a mechanism that alters the heat capacity of an enzyme. We utilised MMRT to investigate the heat capacity of the enzyme MalL and how this could be altered by strategic mutation of specific residues. Robust kinetic data was collected using a specialised Stopped-Flow apparatus and two crystallographic structures were solved during this investigation. Our results are a promising first step in showing how changes in enzyme activity can be linked to changes in enzyme dynamics in auxiliary domains far from the active site using MMRT and heat capacity as a model.

6.1 Future Work

Some inconsistencies were observed in the kinetics of MalL W531G and S536R so these would need to be further investigated and repeated. MalL W531G showed a contrasting result to indole between assays. In both assays performed indole

influenced enzyme activity. This indicates that the system is well suited for further studying the effects of allostery. If the allosteric response in MalL W531G can be optimised other mutations (such as S536R) could potentially be introduced to probe how multiple mutations alter enzyme dynamics in relation to each other. Obtaining a crystal structure for MalL W531G with and without indole bound could give key insights into the mechanisms of allostery. In addition, the W531A mutant could be re-cloned and similarly characterised. The W531A mutant would allow for characterisation of allostery with different binding affinities of indole by introducing a small steric constraint in the binding pocket.

Some work has been done to produce a MalL mutant without the C-terminal domain (unpublished), however this failed to produce soluble protein. Additional constructs are being designed in collaboration with the Institute for Protein Innovation (Harvard) to pursue obtaining soluble constructs. Kinetic and structural characterisation of these mutants would give insights into the function of the C-terminal domain as a whole and provide context to the ability of mutations in this domain to affect the rate.

Molecular dynamics simulations would well complement the current kinetic and structural data collected. Molecular dynamics allows for an atomic level overview of the link between structural changes and enzyme dynamics. This would provide insight into the mechanism by which introduced mutations alter the enzymatic rate.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography, 66*(Pt 2), 213-221.

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., & Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D, 68*(4), 352-367.

Appling, D. R., Anthony-Cahill, S. J., & Mathews, C. K. (2019). *Biochemistry : concepts and connections*. (Second edition, global edition. ed.). New York: Pearson Education Limited.

Arcus, V. L., Prentice, E. J., Hobbs, J. K., Mulholland, A. J., Van der Kamp, M. W., Pudney, C. R., Parker, E. J., & Schipper, L. A. (2016). On the Temperature Dependence of Enzyme-Catalyzed Rates. *Biochemistry, 55*(12), 1681-8.

Benkovic, S. J., & Hammes-Schiffer, S. (2003). A Perspective on Enzyme Catalysis. *Science, 301*(5637), 1196-1202.

Berg, J. M., Tymoczko, J. L., Gatto, G. J., & Stryer, L. (2015). *Biochemistry*. (Eighth edition. ed.). New York: W.H. Freeman & Company, a Macmillan Education Imprint.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1965). Structure of Hen Egg-White Lysozyme: A Three-dimensional Fourier Synthesis at 2 Å Resolution. *Nature, 206*(4986), 757-761.

Blow, D. M. (2002). *Outline of crystallography for biologists*. Oxford ; New York: Oxford University Press.

Chen, W. W., Niepel, M., & Sorger, P. K. (2010). Classic and contemporary approaches to modeling biochemical reactions. *Genes & development, 24*(17), 1861-1875.

Colombo, M., Rau, D., & Parsegian, V. (1992). Protein solvation in allosteric regulation: a water effect on hemoglobin. *Science, 256*(5057), 655-659.

Cooper, A., & Dryden, D. T. F. (1984). Allostery without conformational change. *European Biophysics Journal, 11*(2), 103-109.

Corel Corporation. (2019). Gravit Designer. https://www.designer.io/en/.

Cui, Q., & Karplus, M. (2008). Allostery and cooperativity revisited. *Protein Science, 17*(8), 1295-1307.

Dang Ngoc Chan, C. (2011). *Braggs Law*. Retrieved November, 2019, from https://en.wikipedia.org/wiki/File:Braggs_Law.svg.

Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure, 3*(9), 853-859.

Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schülein, M., & Withers, S. G. (1998). Snapshots along an Enzymatic Reaction Coordinate: Analysis of a Retaining β-Glycoside Hydrolase. *Biochemistry, 37*(34), 11707-11713.

Deckert, K., Budiardjo, S. J., Brunner, L. C., Lovell, S., & Karanicolas, J. (2012). Designing allosteric control into enzymes by chemical rescue of structure. *Journal of the American Chemical Society, 134*(24), 10055-60.

del Sol, A., Tsai, C. J., Ma, B. Y., & Nussinov, R. (2009). The Origin of Allosteric Functional Modulation: Multiple Pre-existing Pathways. *Structure, 17*(8), 1042-1050.

Dodson, E. J. (1991). Baverage.

Drenth, J. (1999). *Principles of protein x-ray crystallography*. Springer advanced texts in chemistry (2nd ed.). New York: Springer.

Eaton, W. A., Henry, E. R., Hofrichter, J., Bettati, S., Viappiani, C., & Mozzarelli, A. (2007). Evolution of allosteric models for hemoglobin. *IUBMB Life, 59*(8-9), 586-599.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D, 66*(4), 486-501.

Evans, M. G. (1938). Thermodynamical treatment of transition state. *Transactions of the Faraday Society, 34*(0), 49-57.

Evans, P. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D, 67*(4), 282-292.

Evans, P. R., & Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography, 69*(Pt 7), 1204-1214.

Feller, G., & Gerday, C. (2003). Psychrophilic enzymes: hot topics in cold adaptation. *Nature Reviews Microbiology, 1*(3), 200-208.

Fersht, A. (1999). *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding*. New York: W.H. Freeman.

Fischer, H. E., & Pinner, A. (1894). Sitzung vom 15. October 1894. *Berichte der deutschen chemischen Gesellschaft, 27*(3), 2643-2658.

Garrett, R., & Grisham, C. M. (2013). *Biochemistry*. (5th ed.). Belmont, CA: Brooks/Cole, Cengage Learning.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571-607). Totowa, NJ: Humana Press.

Gómez, J., Hilser, V. J., Xie, D., & Freire, E. (1995). The heat capacity of proteins. *Proteins: Structure, Function, and Bioinformatics, 22*(4), 404-412.

Haldane, J. B. S. (1930). *Enzymes*. Monographs on biochemistry. London, New York,: Longmans, Green.

Hertig, S., Latorraca, N. R., & Dror, R. O. (2016). Revealing Atomic-Level Mechanisms of Protein Allostery with Molecular Dynamics Simulations. *PLOS Computational Biology, 12*(6), e1004746.

Hilser, V. J., Wrabl, J. O., & Motlagh, H. N. (2012). Structural and Energetic Basis of Allostery. *Annual Review of Biophysics, 41*(1), 585-609.

Hobbs, J. K., Jiao, W., Easter, A. D., Parker, E. J., Schipper, L. A., & Arcus, V. L. (2013). Change in Heat Capacity for Enzyme Catalysis Determines Temperature Dependence of Enzyme Catalyzed Rates. *ACS Chemical Biology, 8*(11), 2388-2393.

Justi, R., & Gilbert, J. K. (1999). History and Philosophy of Science through Models: The Case of Chemical Kinetics. *Science & Education, 8*(3), 287-307.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D, 66*(2), 133-144.

Kantardjieff, K. A., & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein science : a publication of the Protein Society, 12*(9), 1865-1871.

Koshland, D. E., Némethy, G., & Filmer, D. (1966). Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits*. *Biochemistry, 5*(1), 365-385.

Kraakman, K. L. (2017). *Testing Macromolecular Rate Theory.* Master of Science (Research) thesis, University of Waikato, Hamilton, New Zealand.

Le Novere, N. (2008). *Allostery*. Retrieved November, 2019, from https://en.wikipedia.org/wiki/File:Allostery.png.

Lemieux, R. U., & Huber, G. (1955). The Solvolysis of the Alpha- and Beta-3,4,6-Tri-O-Acetyl-D-Glucopyranosyl Chlorides. *Canadian Journal of Chemistry, 33*(1), 128-133.

Liao, J., Okuyama, M., Ishihara, K., Yamori, Y., Iki, S., Tagami, T., Mori, H., Chiba, S., & Kimura, A. (2016). Kinetic properties and substrate inhibition of α-galactosidase from Aspergillus niger. *Bioscience, Biotechnology, and Biochemistry, 80*(9), 1747-1752.

Liu, J., & Nussinov, R. (2016). Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLOS Computational Biology, 12*(6), e1004966.

Maguire, M. E., & Cowan, J. A. (2002). Magnesium chemistry and biochemistry. *Biometals, 15*(3), 203-10.

Marino, K. A., Shang, Y., & Filizola, M. (2018). Insights into the function of opioid receptors from molecular dynamics simulations of available crystal structures. *British Journal of Pharmacology, 175*(14), 2834-2845.

Matthews, B. W. (1968). Solvent content of protein crystals. *Journal of Molecular Biology, 33*(2), 491-497.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography, 40*(Pt 4), 658-674.

McPherson, A., & Gavira, J. A. (2013). Introduction to protein crystallization. *Acta crystallographica. Section F, Structural biology communications, 70*(Pt 1), 2-20.

Microsoft. (1990). Microsoft PowerPoint.

Monod, J., Wyman, J., & Changeux, J.-P. (1965). On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology, 12*(1), 88-118.

Motlagh, H. N., Wrabl, J. O., Li, J., & Hilser, V. J. (2014). The ensemble nature of allostery. *Nature, 508*(7496), 331-339.

Murshudov, G. N., Vagin, A. A., & Dodson, E. J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D, 53*(3), 240-255.

Nicholls, R. A., Fischer, M., McNicholas, S., & Murshudov, G. N. (2014). Conformation-independent structural comparison of macromolecules with ProSMART. *Acta crystallographica. Section D, Biological crystallography, 70*(Pt 9), 2487-2499.

Parker, M. J., Spencer, J., & Clarke, A. R. (1995). An Integrated Kinetic Analysis of Intermediates and Transition States in Protein Folding Reactions. *Journal of Molecular Biology, 253*(5), 771-786.

Pauling, L. (1948). Nature of Forces between Large Molecules of Biological Interest. *Nature, 161*(4097), 707-709.

Perkin Elmer Inc. (1998). ChemDraw Prime.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., & North, A. C. T. (1960). Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature, 185*(4711), 416-422.

Phillips, D. C. (1966). The three-dimensional structure of an enzyme molecule. *Scientific American, 215*(5), 78-90.

Rodrigues, C. H., Pires, D. E., & Ascher, D. B. (2018). DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Research, 46*(W1), W350-W355.

Schipper, L. A., Hobbs, J. K., Rutledge, S., & Arcus, V. L. (2014). Thermodynamic theory explains the temperature optima of soil microbial processes and high Q10 values at low temperatures. *Global Change Biology, 20*(11), 3578-3586.

Schönert, S., Buder, T., & Dahl, M. K. (1998). Identification and Enzymatic Characterization of the Maltose-Inducible α-Glucosidase MalL (Sucrase-Isomaltase-Maltase) of *Bacillus subtilis*. *Journal of Bacteriology, 180*(9), 2574-2578.

Schrödinger, L. (2000). The PyMOL Molecular Graphics System.

Shafee, T. (2015). *Enzyme catalysis energy levels 2*. Retrieved November, 2019, from https://en.wikipedia.org/wiki/File:Enzyme_catalysis_energy_levels_2.svg.

Strynadka, N. C. J., & James, M. N. G. (1991). Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. *Journal of Molecular Biology, 220*(2), 401-424.

Sun, Z., Liu, Q., Qu, G., Feng, Y., & Reetz, M. T. (2019). Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chemical Reviews, 119*(3), 1626-1665.

Thomas, M. T., & Scopes, K. R. (1998). The effects of temperature on the kinetics and stability of mesophilic and thermophilic 3-phosphoglycerate kinases. *Biochemical Journal, 330*(3), 1087-1095.

Tonge, P. J., & Carey, P. R. (1992). Forces, bond lengths, and reactivity: fundamental insight into the mechanism of enzyme catalysis. *Biochemistry, 31*(38), 9122-9125.

Vagin, A., & Teplyakov, A. (1997). MOLREP: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography, 30*(6), 1022-1025.

van der Kamp, M. W., Prentice, E. J., Kraakman, K. L., Connolly, M., Mulholland, A. J., & Arcus, V. L. (2018). Dynamical origins of heat capacity changes in enzyme-catalysed reactions. *Nature Communications, 9*(1), 1177.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D, 67*(4), 235-242.

Wolde, P. R. t., & Frenkel, D. (1997). Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations. *Science, 277*(5334), 1975-1978.

Wong, C., Sridhara, S., Bardwell, J. C. A., & Jakob, U. (2000). Heating Greatly Speeds Coomassie Blue Staining and Destaining. *BioTechniques, 28*(3), 426-432.

Xia, Y., DiPrimio, N., Keppel, T. R., Vo, B., Fraser, K., Battaile, K. P., Egan, C., Bystroff, C., Lovell, S., Weis, D. D., Anderson, J. C., & Karanicolas, J. (2013). The Designability of Protein Switches by Chemical Rescue of Structure: Mechanisms of Inactivation and Reactivation. *Journal of the American Chemical Society, 135*(50), 18840-18849.

7 Appendix

7.1 Gene and protein sequences of MalL variants

Mutation positions for MalL variants are in bolded and underlined.

7.1.1 MalL WT

7.1.1.1 Nucleotide sequence – 1764 bp

ATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTA
TTTTCAGGGCGCCATGGGTGAATGGTGGAAAGAAGCCGTTGTTTATCAGATTTATCCGCGT
AGCTTTTATGATGCCAATGGTGATGGTTTTGGTGATCTGCAAGGTGTTATTCAGAAACTGG
ATTACATCAAAAATCTGGGTGCCGATGTGATTTGGCTGAGTCCGGTTTTTGATAGTCCGCA
GGATGATAATGGTTATGATATTAGCGATTACAAAAACATGTATGAAAAATTTGGCACCAAT
GAAGATATGTTTCAGCTGATTGATGAAGTGCATAAACGCGGTATGAAAATTGTTATGGATC
TGGTGGTTAATCATACCAGTGATGAACATGCATGGTTTGCAGAAAGCCGTAAAAGCAAAG
ATAATCCGTATCGCGATTATTATCTGTGGAAAGATCCGAAACCGGATGGTAGCGAACCGA
ATAATTGGGGTAGCATTTTTAGCGGTAGCGCATGGACCTATGATGAAGGCACCGGTCAGT
ATTATCTGCATTACTTTAGCAAAAAACAGCCGGATCTGAATTGGGAAAATGAAGCAGTTCG
TCGTGAAGTGTATGATGTTATGCGTTTTTGGATGGATCGTGGTGTTGATGGTTGGCGTATG
GATGTTATTGGTAGCATTAGCAAATATACCGATTTTCCGGATTATGAAACCGATCATAGCC
GTAGCTATATTGTGGGTCGTTATCATAGCAATGGTCCGCGTCTGCATGAATTTATTCAGGA
AATGAATCGCGAAGTTCTGAGCCATTATGATTGTATGACCGTTGGTGAAGCAAATGGCAG
CGATATTGAAGAAGCCAAAAAATACACAGATGCCAGTCGCCAAGAACTGAATATGATTTTT
ACCTTTGAACATATGGATATTGATAAAGAACAGAATAGCCCGAATGGCAAATGGCAGATT
AAACCGTTTGATCTGATTGCCCTGAAAAAAACCATGACCCGTTGGCAGACAGGTCTGATGA
ATGTTGGTTGGAATACCCTGTATTTTGAAAATCATGATCAGCCTCGTGTTATTAGCCGTTGG
GGTAATGATCGTAAACTGCGTAAAGAATGTGCAAAAGCATTTGCAACCGTTCTGCATGGTA
TGAAAGGCACCCCGTTTATCTATCAGGGTGAAGAAATTGGTATGGTGAATAGCGATATGC
CGCTGGAAATGTATGATGATCTGGAAATCAAAAATGCCTATCGCGAACTGGTGGTGGAAA
ACAAAACCATGAGCGAAAAAGAATTTGTGAAAGCCGTGATGATTAAAGGTCGTGATCATG
CACGTACCCCGATGCAGTGGGATGCAGGTAAACATGCAGGTTTTACCGCAGGCGATCCGT
GGATTCCGGTTAATAGCCGTTATCAGGATATTAATGTGAAAGAAAGCCTGGAAGATCAGG

ATAGCATTTTCTTTTACTATCAGAAACTGATTCAGCTGCGCAAACAGTATAAAATCATGATT
TATGGCGATTATCAGCTGCTGCAAGAAAATGATCCGCAGGTTTTTAGCTATCTGCGTGAAT
ATCGTGGTGAAAAACTGCTGGTTGTTGTTAATCTGAGCGAAGAAAAAGCACTGTTTGAAG
CACCTCCGGAACTGATTCATGAACGTTGGAAAGTTCTGATTAGCAATTATCCGCAGGAACG
TGCAGATCTGAAAAGCATTAGCCTGAAACCGTATGAAGCAGTTATGGGCATTAGCATCTAA
TAA

7.1.1.2 Protein sequence – 586 amino acids – 69.15 kDa

MSYYHHHHHHDYDIPTTENLYFQGAMGEWWKEAVVYQIYPRSFYDANGDGFGDLQGVIQK
LDYIKNLGADVIWLSPVFDSPQDDNGYDISDYKNMYEKFGTNEDMFQLIDEVHKRGMKIVMD
LVVNHTSDEHAWFAESRKSKDNPYRDYYLWKDPKPDGSEPNNWGSIFSGSAWTYDEGTGQY
YLHYFSKKQPDLNWENEAVRREVYDVMRFWMDRGVDGWRMDVIGSISKYTDFPDYETDHS
RSYIVGRYHSNGPRLHEFIQEMNREVLSHYDCMTVGEANGSDIEEAKKYTDASRQELNMIFTFE
HMDIDKEQNSPNGKWQIKPFDLIALKKTMTRWQTGLMNVGWNTLYFENHDQPRVISRWG
NDRKLRKECAKAFATVLHGMKGTPFIYQGEEIGMVNSDMPLEMYDDLEIKNAYRELVVENKT
MSEKEFVKAVMIKGRDHARTPMQWDAGKHAGFTAGDPWIPVNSRYQDINVKESLEDQDSIF
FYYQKLIQLRKQYKIMIYGDYQLLQENDPQVFSYLREYRGEKLLVVVNLSEEKALFEAPPELIHER
WKVLISNYPQERADLKSISLKPYEAVMGISI*

7.1.2 MalL S536R

7.1.2.1 Nucleotide sequence -1764 bp

ATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTA
TTTTCAGGGCGCCATGGGTGAATGGTGGAAAGAAGCCGTTGTTTATCAGATTTATCCGCGT
AGCTTTTATGATGCCAATGGTGATGGTTTTGGTGATCTGCAAGGTGTTATTCAGAAACTGG
ATTACATCAAAAATCTGGGTGCCGATGTGATTTGGCTGAGTCCGGTTTTTGATAGTCCGCA
GGATGATAATGGTTATGATATTAGCGATTACAAAAACATGTATGAAAAATTTGGCACCAAT
GAAGATATGTTTCAGCTGATTGATGAAGTGCATAAACGCGGTATGAAAATTGTTATGGATC
TGGTGGTTAATCATACCAGTGATGAACATGCATGGTTTGCAGAAAGCCGTAAAAGCAAAG
ATAATCCGTATCGCGATTATTATCTGTGGAAAGATCCGAAACCGGATGGTAGCGAACCGA
ATAATTGGGGTAGCATTTTTAGCGGTAGCGCATGGACCTATGATGAAGGCACCGGTCAGT
ATTATCTGCATTACTTTAGCAAAAAACAGCCGGATCTGAATTGGGAAAATGAAGCAGTTCG

TCGTGAAGTGTATGATGTTATGCGTTTTTGGATGGATCGTGGTGTTGATGGTTGGCGTATG
GATGTTATTGGTAGCATTAGCAAATATACCGATTTTCCGGATTATGAAACCGATCATAGCC
GTAGCTATATTGTGGGTCGTTATCATAGCAATGGTCCGCGTCTGCATGAATTTATTCAGGA
AATGAATCGCGAAGTTCTGAGCCATTATGATTGTATGACCGTTGGTGAAGCAAATGGCAG
CGATATTGAAGAAGCCAAAAAATACACAGATGCCAGTCGCCAAGAACTGAATATGATTTTT
ACCTTTGAACATATGGATATTGATAAAGAACAGAATAGCCCGAATGGCAAATGGCAGATT
AAACCGTTTGATCTGATTGCCCTGAAAAAAACCATGACCCGTTGGCAGACAGGTCTGATGA
ATGTTGGTTGGAATACCCTGTATTTTGAAAATCATGATCAGCCTCGTGTTATTAGCCGTTGG
GGTAATGATCGTAAACTGCGTAAAGAATGTGCAAAAGCATTTGCAACCGTTCTGCATGGTA
TGAAAGGCACCCCGTTTATCTATCAGGGTGAAGAAATTGGTATGGTGAATAGCGATATGC
CGCTGGAAATGTATGATGATCTGGAAATCAAAAATGCCTATCGCGAACTGGTGGTGGAAA
ACAAAACCATGAGCGAAAAAGAATTTGTGAAAGCCGTGATGATTAAAGGTCGTGATCATG
CACGTACCCCGATGCAGTGGGATGCAGGTAAACATGCAGGTTTTACCGCAGGCGATCCGT
GGATTCCGGTTAATAGCCGTTATCAGGATATTAATGTGAAAGAAAGCCTGGAAGATCAGG
ATAGCATTTTCTTTTACTATCAGAAACTGATTCAGCTGCGCAAACAGTATAAAATCATGATT
TATGGCGATTATCAGCTGCTGCAAGAAAATGATCCGCAGGTTTTTAGCTATCTGCGTGAAT
ATCGTGGTGAAAAACTGCTGGTTGTTGTTAATCTGAGCGAAGAAAAAGCACTGTTTGAAG
CACCTCCGGAACTGATTCATGAACGTTGGAAAGTTCTGATT**CGC**AATTATCCGCAGGAACG
TGCAGATCTGAAAAGCATTAGCCTGAAACCGTATGAAGCAGTTATGGGCATTAGCATCTAA
TAA

7.1.2.2 Protein sequence – 586 amino acids – 69.22 kDa

MSYYHHHHHHDYDIPTTENLYFQGAMGEWWKEAVVYQIYPRSFYDANGDGFGDLQGVIQK
LDYIKNLGADVIWLSPVFDSPQDDNGYDISDYKNMYEKFGTNEDMFQLIDEVHKRGMKIVMD
LVVNHTSDEHAWFAESRKSKDNPYRDYYLWKDPKPDGSEPNNWGSIFSGSAWTYDEGTGQY
YLHYFSKKQPDLNWENEAVRREVYDVMRFWMDRGVDGWRMDVIGSISKYTDFPDYETDHS
RSYIVGRYHSNGPRLHEFIQEMNREVLSHYDCMTVGEANGSDIEEAKKYTDASRQELNMIFTFE
HMDIDKEQNSPNGKWQIKPFDLIALKKTMTRWQTGLMNVGWNTLYFENHDQPRVISRWG
NDRKLRKECAKAFATVLHGMKGTPFIYQGEEIGMVNSDMPLEMYDDLEIKNAYRELVVENKT
MSEKEFVKAVMIKGRDHARTPMQWDAGKHAGFTAGDPWIPVNSRYQDINVKESLEDQDSIF
FYYQKLIQLRKQYKIMIYGDYQLLQENDPQVFSYLREYRGEKLLVVVNLSEEKALFEAPPELIHER
WKVLI**R**NYPQERADLKSISLKPYEAVMGISI*

7.1.3 MalL TM

7.1.3.1 Nucleotide sequence (E. coli optimised) – 1764 bp

ATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTA
TTTTCAGGGCGCCATGGGTGAATGGTGGAAAGAAGCAGTGGTTTATCAGATTTATCCGCG
TAGCTTTTATGATGCCAATGGTGATGGTTTTGGTGATCTGCAAGGTGTTATTCAGAAACTG
GACTACATTAAAAACCTGGGTGCAGATGTTATTTGGCTGAGTCCGGTTTTTGATAGTCCGC
AGGATGATAATGGTTATGATATCAGCGACTATAAAAACATGTACGAGAAATTTGGCACCAA
CGAGGATATGTTTCAGCTGATTGATGAAGTGCATAAACGCGGTATGAAAATCGTTATGGAT
CTGGTTGTTAATCACACCAGTGATGAACATGCATGGTTTGCAGAAAGCCGTAAAAGCAAA
GATAATCCGTATCGCGATTACTACCTGTGGAAAGATCCGAAACCGGATGGTAGCGAACCG
AATAATTGGGGTAGCATTTTTAGCGGTAGCGCATGGACCTATGATGAAGGCACCGGTCAG
TATTATCTGCACTATTTTAGCAAAAAACAGCCGGATCTGAACTGGGAAAATGAAGCAGTTC
GTCGTGAAGTTTATGATGTGATGCGTTTTTGGATGGATCGTGGTGTTGATGGTTGGCGTAT
GGATGTTATTGGTAGCATTAGCAAATATACCGACTTTCCGGATTATGAAACCGATCATAGC
CGTAGCTATATTGTGGGTCGTTATCATAGCAATGGTCCGCGTCTGCATGAGTTTATTCAAG
AAATGAATCGTGAAGTGCTGAGCCACTATGATTGTATGACCGTTGGTGAAGCAAATGGCA
GCGATATTGAAGAAGCCAAAAAGTATACCGATGCAAGCCGTCAAGAACTGAACATGATTT
TTACCTTTGAGCACATGGACATCGATAAAGAACAGAATAGCCCGAATGGTAAATGGCAGA
TTAAACCGTTTGATCTGATCGCCCTGAAAAAAACCATGACACGTTGGCAGACAGGTCTGAT
GAATGTTGGTTGGAATACCCTGTATTTTGAGAACCATGATCAGCCTCGTGTTATTAGCCGTT
GGGGTAATGATCGTAAACTGCGTAAAGAATGTGCAAAAGCATTTGCAACCGTTCTGCATG
GTATGAAAGGCACCCCGTTTATCTATCAGGGTGAAGAAATTGGTATGGTGAATAGCGATA
TGCCGCTGGAAATGTATGATGACCTGGAAATCAAAAATGCCTATCGTGAACTGGTGGTGG
AAAACAAAACCATGTCCGAAAAAGAATTTGTGAAGGCCGTGATGATTAAAGGTCGTGATC
ATGCACGTACCCCGATGCAGTGGGATGCAGGTAAACATGCCGGTTTTACCGCAGGCGATC
CGTGGATTCCGGTTAATAGCCGTTATCAGGATATCAACGTTAAAGAGAGCCTGGAAGATC
AGGATAGCATCTTCTTCTATTACCAGAAACTGATTCAGCTGCGCAAACAGTACAAAATCAT
GATCTATGGTGATTATCAGCTGCTGCAAGAAAATGATCCGCAGGTTTTTTCATATCTGCGTG
AATATCGTGGTGAGAAACTGCTGGTTGTTGTGAATCTGAGCGAAGAAAAAGCACTGTTTG
AAGCACCGCCTGAACTGATCCATGAACGCTGGAAAGTTCTGATT**GAT**AACTATCCGCAAGA

ACGTGCAGATCTGAAAAGCATTAGCCTGAAACCGTAT**CAG**GCA**CGT**ATGGGTATTAGCAT
CTAATAA

7.1.3.2 Protein sequence – 586 amino acids – 69.23 kDa

MSYYHHHHHHDYDIPTTENLYFQGAMGEWWKEAVVYQIYPRSFYDANGDGFGDLQGVIQK
LDYIKNLGADVIWLSPVFDSPQDDNGYDISDYKNMYEKFGTNEDMFQLIDEVHKRGMKIVMD
LVVNHTSDEHAWFAESRKSKDNPYRDYYLWKDPKPDGSEPNNWGSIFSGSAWTYDEGTGQY
YLHYFSKKQPDLNWENEAVRREVYDVMRFWMDRGVDGWRMDVIGSISKYTDFPDYETDHS
RSYIVGRYHSNGPRLHEFIQEMNREVLSHYDCMTVGEANGSDIEEAKKYTDASRQELNMIFTFE
HMDIDKEQNSPNGKWQIKPFDLIALKKTMTRWQTGLMNVGWNTLYFENHDQPRVISRWG
NDRKLRKECAKAFATVLHGMKGTPFIYQGEEIGMVNSDMPLEMYDDLEIKNAYRELVVENKT
MSEKEFVKAVMIKGRDHARTPMQWDAGKHAGFTAGDPWIPVNSRYQDINVKESLEDQDSIF
FYYQKLIQLRKQYKIMIYGDYQLLQENDPQVFSYLREYRGEKLLVVVNLSEEKALFEAPPELIHER
WKVLI**D**NYPQERADLKSISLKPY**Q**A**R**MGISI*

7.1.4 MalL W531G

7.1.4.1 Nucleotide sequence – 1764 bp

ATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTA
TTTTCAGGGCGCCATGGGTGAATGGTGGAAAGAAGCCGTTGTTTATCAGATTTATCCGCGT
AGCTTTTATGATGCCAATGGTGATGGTTTTGGTGATCTGCAAGGTGTTATTCAGAAACTGG
ATTACATCAAAAATCTGGGTGCCGATGTGATTTGGCTGAGTCCGGTTTTTGATAGTCCGCA
GGATGATAATGGTTATGATATTAGCGATTACAAAAACATGTATGAAAAATTTGGCACCAAT
GAAGATATGTTTCAGCTGATTGATGAAGTGCATAAACGCGGTATGAAAATTGTTATGGATC
TGGTGGTTAATCATACCAGTGATGAACATGCATGGTTTGCAGAAAGCCGTAAAAGCAAAG
ATAATCCGTATCGCGATTATTATCTGTGGAAAGATCCGAAACCGGATGGTAGCGAACCGA
ATAATTGGGGTAGCATTTTTAGCGGTAGCGCATGGACCTATGATGAAGGCACCGGTCAGT
ATTATCTGCATTACTTTAGCAAAAAACAGCCGGATCTGAATTGGGAAAATGAAGCAGTTCG
TCGTGAAGTGTATGATGTTATGCGTTTTTGGATGGATCGTGGTGTTGATGGTTGGCGTATG
GATGTTATTGGTAGCATTAGCAAATATACCGATTTTCCGGATTATGAAACCGATCATAGCC
GTAGCTATATTGTGGGTCGTTATCATAGCAATGGTCCGCGTCTGCATGAATTTATTCAGGA
AATGAATCGCGAAGTTCTGAGCCATTATGATTGTATGACCGTTGGTGAAGCAAATGGCAG

CGATATTGAAGAAGCCAAAAAATACACAGATGCCAGTCGCCAAGAACTGAATATGATTTTT
ACCTTTGAACATATGGATATTGATAAAGAACAGAATAGCCCGAATGGCAAATGGCAGATT
AAACCGTTTGATCTGATTGCCCTGAAAAAAACCATGACCCGTTGGCAGACAGGTCTGATGA
ATGTTGGTTGGAATACCCTGTATTTTGAAAATCATGATCAGCCTCGTGTTATTAGCCGTTGG
GGTAATGATCGTAAACTGCGTAAAGAATGTGCAAAAGCATTTGCAACCGTTCTGCATGGTA
TGAAAGGCACCCCGTTTATCTATCAGGGTGAAGAAATTGGTATGGTGAATAGCGATATGC
CGCTGGAAATGTATGATGATCTGGAAATCAAAAATGCCTATCGCGAACTGGTGGTGGAAA
ACAAAACCATGAGCGAAAAAGAATTTGTGAAAGCCGTGATGATTAAAGGTCGTGATCATG
CACGTACCCCGATGCAGTGGGATGCAGGTAAACATGCAGGTTTTACCGCAGGCGATCCGT
GGATTCCGGTTAATAGCCGTTATCAGGATATTAATGTGAAAGAAAGCCTGGAAGATCAGG
ATAGCATTTTCTTTTACTATCAGAAACTGATTCAGCTGCGCAAACAGTATAAAATCATGATT
TATGGCGATTATCAGCTGCTGCAAGAAAATGATCCGCAGGTTTTTAGCTATCTGCGTGAAT
ATCGTGGTGAAAAACTGCTGGTTGTTGTTAATCTGAGCGAAGAAAAAGCACTGTTTGAAG
CACCTCCGGAACTGATTCATGAACGT**GGG**AAAGTTCTGATTAGCAATTATCCGCAGGAACG
TGCAGATCTGAAAAGCATTAGCCTGAAACCGTATGAAGCAGTTATGGGCATTAGCATCTAA
TAA

7.1.4.2 Protein sequence – 586 amino acids – 69.02 kDa

MSYYHHHHHHDYDIPTTENLYFQGAMGEWWKEAVVYQIYPRSFYDANGDGFGDLQGVIQK
LDYIKNLGADVIWLSPVFDSPQDDNGYDISDYKNMYEKFGTNEDMFQLIDEVHKRGMKIVMD
LVVNHTSDEHAWFAESRKSKDNPYRDYYLWKDPKPDGSEPNNWGSIFSGSAWTYDEGTGQY
YLHYFSKKQPDLNWENEAVRREVYDVMRFWMDRGVDGWRMDVIGSISKYTDFPDYETDHS
RSYIVGRYHSNGPRLHEFIQEMNREVLSHYDCMTVGEANGSDIEEAKKYTDASRQELNMIFTFE
HMDIDKEQNSPNGKWQIKPFDLIALKKTMTRWQTGLMNVGWNTLYFENHDQPRVISRWG
NDRKLRKECAKAFATVLHGMKGTPFIYQGEEIGMVNSDMPLEMYDDLEIKNAYRELVVENKT
MSEKEFVKAVMIKGRDHARTPMQWDAGKHAGFTAGDPWIPVNSRYQDINVKESLEDQDSIF
FYYQKLIQLRKQYKIMIYGDYQLLQENDPQVFSYLREYRGEKLLVVVNLSEEKALFEAPPELIHER
GKVLISNYPQERADLKSISLKPYEAVMGISI*

7.2 Example analysis of kinetic data

Initial kinetic data (V_{max}) is collected as change in absorbance per second by linear regression of the first 10 seconds of rate data (Table 7.1). The V_{max} is converted to M sec^{-1} by dividing by the extinction coefficient for the enzyme at the relevant pH and wavelength (13270 M^{-1}) (Kraakman, 2017) (Table 7.2). The k_{cat} is then calculated by dividing by the enzyme concentration (1.87 X 10^{-7} M) (Table 7.3). Temperature data is plotted as the natural log of k_{cat} versus temperature. This data can be seen in Section 3 (Figure 3.11).

Table 7.1.Example linear regression of spectroscopy data

		V_{max} (ΔAbs sec^{-1})			
	Temperature (K)	**1**	**2**	**3**	**4**
MalL WT	**288.0**	0.01446	0.01515	0.01467	0.0147
	292.7	0.02773	0.02902	0.02873	0.02882
	297.7	0.05556	0.05555	0.05519	0.0549
	302.6	0.09339	0.09609	0.09514	0.09515
	305.9		0.13791	0.14062	0.14084
	307.6		0.17075	0.17022	0.17012
	309.7		0.19833	0.19563	0.19771
	311.9		0.20672	0.20792	0.20713
	313.8		0.20159	0.20019	0.2001
	315.9		0.21916	0.22114	0.22076
	317.8		0.19737	0.20065	0.19952
	319.8		0.15637	0.15598	0.15573
MalL S536R	**288.1**	0.00954	0.00978	0.00969	0.00982
	292.8	0.01808	0.0185	0.01834	0.01815
	297.6	0.0331	0.03257	0.0329	0.03265
	302.7	0.05533	0.05479	0.05555	0.05507
	307.4	0.08154	0.08576	0.08668	0.08647
	311.8		0.11055	0.11214	0.11257
	313.9	0.11501	0.11717	0.11748	0.11682
	315.9	0.11118	0.1158	0.11552	0.11574
	317.9		0.12368	0.12442	0.12346
	319.9	0.11303	0.11673	0.11665	0.11644
	321.8	0.10125	0.09957	0.09935	0.09949
	323.7	0.08189	0.07852	0.07874	0.08006

Table 7.2. Example calculation of V_{max}

		V_{max} (M sec^{-1})			
	Temperature (K)	**1**	**2**	**3**	**4**
MalL WT	**288.1**	1.08968E-06	1.14167E-06	1.1055E-06	1.10776E-06
	292.8	2.08968E-06	2.18689E-06	2.16503E-06	2.17182E-06
	297.6	4.18689E-06	4.18613E-06	4.15901E-06	4.13715E-06
	302.7	7.03768E-06	7.24115E-06	7.16956E-06	7.17031E-06
	307.4		1.03926E-05	1.05968E-05	1.06134E-05
	311.8		1.28674E-05	1.28274E-05	1.28199E-05
	313.9		1.49457E-05	1.47423E-05	1.4899E-05
	315.9		1.5578E-05	1.56684E-05	1.56089E-05
	317.9		1.51914E-05	1.50859E-05	1.50791E-05
	319.9		1.65154E-05	1.66647E-05	1.6636E-05
	321.8		1.48734E-05	1.51206E-05	1.50354E-05
	323.7		1.17837E-05	1.17543E-05	1.17355E-05
MalL S536R	**288.1**	7.18915E-07	7.37001E-07	7.30219E-07	7.40015E-07
	292.8	1.36247E-06	1.39412E-06	1.38206E-06	1.36775E-06
	297.6	2.49435E-06	2.45441E-06	2.47928E-06	2.46044E-06
	302.7	4.16956E-06	4.12886E-06	4.18613E-06	4.14996E-06
	307.4	6.14469E-06	6.4627E-06	6.53203E-06	6.5162E-06
	311.8		8.33082E-06	8.45064E-06	8.48304E-06
	313.9	8.66692E-06	8.82969E-06	8.85305E-06	8.80332E-06
	315.9	8.3783E-06	8.72645E-06	8.70535E-06	8.72193E-06
	317.9		9.32027E-06	9.37604E-06	9.30369E-06
	319.9	8.51771E-06	8.79653E-06	8.7905E-06	8.77468E-06
	321.8	7.62999E-06	7.50339E-06	7.48681E-06	7.49736E-06
	323.7	6.17106E-06	5.91711E-06	5.93369E-06	6.03316E-06

Table 7.3. Example calculation of k_{cat}

		k_{cat} (sec^{-1})			
	Temperature (K)	**1**	**2**	**3**	**4**
MalL WT	**288.1**	5.82714	6.10520	5.91177	5.92386
	292.8	11.17474	11.69459	11.57772	11.61399
	297.6	22.38977	22.38574	22.24067	22.12380
	302.7	37.63465	38.72270	38.33987	38.34390
	307.4		55.57548	56.66757	56.75622
	311.8		68.80947	68.59588	68.55559
	313.9		79.92376	78.83570	79.67391
	315.9		83.30479	83.78837	83.47001
	317.9		81.23748	80.67331	80.63704
	319.9		88.31791	89.11581	88.96268
	321.8		79.53689	80.85868	80.40331
	323.7		63.01456	62.85740	62.75665
MalL S536R	**288.1**	3.84446	3.94118	3.90491	3.95730
	292.8	7.28595	7.45520	7.39072	7.31415
	297.6	13.33876	13.12518	13.25816	13.15742
	302.7	22.29709	22.07948	22.38574	22.19231
	307.4	32.85929	34.55988	34.93063	34.84600
	311.8		44.54985	45.19059	45.36387
	313.9	46.34715	47.21760	47.34252	47.07655
	315.9	44.80373	46.66551	46.55268	46.64133
	317.9		49.84102	50.13923	49.75237
	319.9	45.54925	47.04029	47.00805	46.92342
	321.8	40.80210	40.12509	40.03643	40.09285
	323.7	33.00033	31.64228	31.73094	32.26287

7.3 Summary kinetic data

7.3.1 Michaelis-Menten kinetics

Table 7.4. Summary Michaelis-Menten constants for MalL Variants

	Enzyme					
Parameter	**MalL WT**	**MalL S536R**	**MalL TM**	**MalL W531G (0mM Indole)**	**MalL W531G (0.5 mM Indole)**	**MalL W531G (5 mM Indole)**
k_{cat} (sec^{-1})	27.1 ± 0.6	25.5 ± 0.5	34.3 ± 1.0	26.1 ± 0.4	30.6 ± 0.4	25.7 ± 0.2
K_M (mM)	0.0817 ± 0.005	0.101 ± 0.006	0.124 ± 0.009	0.0911 ± 0.004	0.0971 ± 0.004	0.123 ± 0.002
K_i (mM)	11.5 ± 1.8	8.7 ± 1.0	3.4 ± 0.3	11.7 ± 1.2	10.7 ± 0.9	26.5 ± 2.4

7.3.2 Temperature characterisation

Table 7.5. Summary MMRT fit parameters of MalL Variants

	Enzyme				
Parameter	**MalL WT**	**MalL S536R**	**MalL TM**	**MalL W531G**	**0.1 M GuHCl**
T_0 (K)	309.4	312.8	309.1	310.6	305.7
$\Delta H^{\ddagger}$ (T_0) (kJ mol^{-1})	42.9 ± 1.2	21.4 ± 0.8	44.9 ± 0.4	39.3 ± 0.5	38.3 ± 1.0
$\Delta S^{\ddagger}$ (T_0) (J mol^{-1} K^{-1})	-70.7 ± 3.7	-145.1 ± 2.5	-62.1 ± 1.4	-81.4 ± 1.5	-87.7 ± 3.4
$\Delta C_P^{\ddagger}$ (T_0) (kJ mol^{-1} K^{-1})	-8.5 ± 0.3	-7.4 ± 0.2	-9.3 ± 0.1	-9.6 ± 0.1	-8.5 ± 0.2
A (kJ mol^{-1} K^{-2})	-218 ± 15	-139 ± 9	-229 ± 6	-221 ± 6	-211 ± 13
T_{opt} (K)	313.4	316.8	313.1	314.6	309.7